

177

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: The Trustees of Columbia University in the City of New York City
- (ii) TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 45
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Cooper & Dunham LLP
(B) STREET: 1185 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: White, John P.
(B) REGISTRATION NUMBER: 28,578
(C) REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPW/MS
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 278-0400
(B) TELEFAX: (212) 391-0525

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20710 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGAGTCGGA GAGTTGGCAC AGGCCTTGAG CTCGCTGTGA CGTTCTCAGG GTGTTGGTTG	60
GGATCAGCTG GTGACTCAGA CAAGTCTTGA GCTCTACAAC GTAACATACG GGCTGATGCC	120
CACCCGATAC CAGAATTACG CAGTCGGCAA TTCTGTGCCC TAGAGTCACC TCAAAGAATA	180
ATCTGTGGTG TCCAAGGGGA GGGTTCTGGG GCCGGCTACT TAGAAACCGC CATAGATCGG	240

178

GCAGGGTGGA GTACTTGAGG AGCCGGCGGT AGGTGGCCAG GTGGGCCCCG TTACCTGCTC 300
TTTTGCGTGC TGCTGGAAGC CTGCTCAGGG ATTTCTTAAC CTCGGCCTCG GTTGGACGTA 360
CCATGGCAGA AGGCGGTTTT GGAGCGGACT CGGTGGGGCG CGGCGGAGAA AAGGCCTCTG 420
TGACTAGGGG AGGCAGGTGG GACTTGGGGA GCTCGGACGA CGAATCAAGC ACCTCCACAA 480
CCAGCACGGA TATGGACGAC CTCCCTGAGG AGAGGAAACC ACTAACGGGA AAGTCTGTAA 540
AAACCTCGTA CATATACGAC GTGCCCACCG TCCCGACCAG CAAGCCGTGG CATTTAATGC 600
ACGACAACTC CCTCTACGCA AGGCCTAGGT TTCCGCCCAG ACCTCTCATA CGGCACCCTT 660
CCGAAAAAGG CAGCATTTTT GCCAGTCGGT TGTCAGCGAC TGACGACGAC TCGGGAGACT 720
ACGCGCCAAT GGATCGCTTC GCCTTCGAGA GCCCCAGGGT GTGTGGTCGC CCTCCCCCTC 780
CGCCTCCAAA TCACCCACCT CCGGCAACTA GGCCGGCAGA CGCGTCAATG GGGGACGTGG 840
GCTGGGCGGA TCTGCAGGGA CTCAAGAGGA CCCCAAAGGG ATTTTTAAAA ACATCTACCA 900
AGGGGGGCAG TCTCAAAGCC CGTGGACGCG ATGTAGGTGA CCGTCTCAGG GACGGCGGCT 960
TTGCCTTTAG TCCTAGGGGC GTGAAATCTG CCATAGGGCA AAACATTAAA TCATGGTTGG 1020
GGATCGGAGA ATCATCGGCG ACTGCTGTCC CCGTCACCAC GCAGCTTATG GTACCGGTGC 1080
ACCTCATTAG AACGCCTGTG ACCGTGGACT ACAGGAATGT TTATTTGCTT TACTTAGAGG 1140
GGGTAATGGG TGTGGGCAA TCAACGCTGG TCAACGCGT GTGCGGGATC TTGCCCCAGG 1200
AGAGAGTGAC AAGTTTTCCC GAGCCCATGG TGTACTGGAC GAGGGCATT ACAGATTGTT 1260
ACAAGGAAAT TTCCCACCTG ATGAAGTCTG GTAAGGCGGG AGACCCGCTG ACGTCTGCCA 1320
AAATATACTC ATGCCAAAAC AAGTTTTCGC TCCCCTTCCG GACGAACGCC ACCGCTATCC 1380
TGCGAATGAT GCAGCCCTGG AACGTTGGGG GTGGGTCTGG GAGGGGCACT CACTGGTGCG 1440
TCTTTGATAG GCATCTCCTC TCCCCAGCAG TGGTGTTCCT TCTCTGCAC CTGAAGCACG 1500
GCCGCCTATC TTTTGATCAC TTCTTTCAAT TACTTTCCAT CTTTAGAGCC ACAGAAGGCG 1560
ACGTGGTCGC CATTCTCACC CTCTCCAGCG CCGAGTCGTT GCGGCGGCTC AGGGCGAGGG 1620
GAAGAAAGAA CGACGGGACG GTGGAGCAA ACTACATCAG AGAATTGGCG TGGGCTTATC 1680
ACGCCGTGTA CTGTTTCATG ATCATGTTGC AGTACATCAC TGTGGAGCAG ATGGTACAAC 1740
TATGCGTACA AACCACAAAT ATTCCGAAA TCTGCTTCCG CAGCGTGCGC CTGGCACACA 1800
AGGAGGAAAC TTTGAAAAAC CTTACGAGC AGAGCATGCT ACCTATGATC ACCCGTGTAC 1860
TGGATCCCGT GAGACATCAT CCCGTCGTGA TCGAGCTTTG CTTTTGTTT TTCACAGAGC 1920
TGAGAAAAAT ACAATTTATC GTAGCCGACG CGGATAAGTT CCACGACGAC GTATGCGGCC 1980
TGTGGACCGA AATCTACAGG CAGATCCTGT CCAATCCGGC TATTAAACCC AGGGCCATCA 2040
ACTGGCCAGC ATTAGAGAGC CAGTCTAAAG CAGTTAATCA CCTAGAGGAG ACATGCAGGG 2100
TCTAGCCTTC TTGGCGGCCC TTGCATGCTG GCGATGCATA TCGTTGACAT GTGGAGCCAC 2160
TGGCGCGTTG CCGACAACGG CGACGACAAT AACCCGCTCC GCCACGCAGC TCATCAATGG 2220
GAGAACCAAC CTCTCCATAG AACTGGAATT CAACGGCACT AGTTTTTTTC TAAATTGGCA 2280

00607479-052000

006290: 6274960

AAATCTGTTG	AATGTGATCA	CGGAGCCGGC	CCTGACAGAG	TTGTGGACCT	CCGCCGAAGT	2340
CGCCGAGGAC	CTCAGGGTAA	CTCTGAAAAA	GAGGCAAAGT	CTTTTTTTCC	CCAACAAGAC	2400
AGTTGTGATC	TCTGGAGACG	GCCATCGCTA	TACGTGCGAG	GTGCCGACGT	CGTCGCAAAC	2460
TTATAACATC	ACCAAGGGCT	TTAACTATAG	CGCTCTGCCC	GGGCACCTTG	GCGGATTTGG	2520
GATCAACGCG	CGTCTGGTAC	TGGGTGATAT	CTTCGCATCA	AAATGGTCGC	TATTCGCGAG	2580
GGACACCCCA	GAGTATCGGG	TGTTTTACCC	AATGAATGTC	ATGGCCGTCA	AGTTTTCCAT	2640
ATCCATTGGC	AACAACGAGT	CGGGCGTAGC	GCTCTATGGA	GTGGTGTCCG	AAGATTTTCGT	2700
GGTCGTCACG	CTCCACAACA	GGTCCAAAGA	GGCTAACGAG	ACGGCGTCCC	ATCTTCTGTT	2760
CGGTCTCCCG	GATTCACTGC	CATCTCTGAA	GGGCCATGCC	ACCTATGATG	AACTCACGTT	2820
CGCCCGAAAC	GCAAAATATG	CGCTAGTGGC	GATCCTGCCT	AAAGATTCTT	ACCAGACACT	2880
CCTTACAGAG	AATTACACTC	GCATATTTCT	GAACATGACG	GAGTCGACGC	CCCTCGAGTT	2940
CACGCGGACG	ATCCAGACCA	GGATCGTATC	AATCGAGGCC	AGGCGCGCCT	GCGCAGCTCA	3000
AGAGGCGGCG	CCGGACATAT	TCTTGGTGTT	TTTCAGATG	TTGGTGGCAC	ACTTTCTTGT	3060
TGCGCGGGGC	ATTGCCGAGC	ACCGATTTGT	GGAGGTGGAC	TGCGTGTGTC	GGCAGTATGC	3120
GGAAGTGTAT	TTTCTCCGCC	GCATCTCGCG	TCTGTGCATG	CCCACGTTCA	CCACTGTCGG	3180
GTATAACCAC	ACCACCCTTG	GCGCTGTGGC	CGCCACACAA	ATAGCTCGCG	TGTCCGCCAC	3240
GAAGTTGGCC	AGTTTGCCCC	GCTCTTCCCA	GGAAACAGTG	CTGGCCATGG	TCCAGCTTGG	3300
CGCCCGTGAT	GGCGCCGTCC	CTTCCTCCAT	TCTGGAGGSC	ATTGCTATGG	TCGTGCAACA	3360
TATGTATACC	GCCTACACTT	ATGTGTACAC	ACTCGGCGAT	ACTGAAAGAA	AATTAATGTT	3420
GGACATACAC	ACGGTCCTCA	CCGACAGCTG	CCCCCCCCAA	GACTCCGGAG	TATCAGAAAA	3480
GCTACTGAGA	ACATATTTGA	TGTTACATC	AATGTGTACC	AACATAGAGC	TGGGCGAAAT	3540
GATCGCCCGC	TTTTCCAAAC	CGGACAGCCT	TAACATCTAT	AGGGGATTCT	CCCCCTGCTT	3600
TCTAGGACTA	AGGTACGATT	TGCATCCAGC	CAAGTTGCGC	GCCGAGGCGC	CGCAGTCGTC	3660
CGCTCTGACG	CGGACTGCCG	TTGCCAGAGG	AACATCGGGA	TTCGCAGAA	TGCTCCACGC	3720
GCTGCACCTC	GATAGCTTAA	ATTTAATTCC	GGCGATTAAC	TGTTCAAAGA	TTACAGCCGA	3780
CAAGATAATA	GCTACGGTAC	CCTTGCCCTCA	CGTCACGTAT	ATCATCAGTT	CCGAAGCACT	3840
CTCGAACGCT	GTTGTCTACG	AGGTGTCGGA	GATCTTCCTC	AAGAGTGCCA	TGTTTATATC	3900
TGCTATCAAA	CCCGATTGCT	CCGGCTTTAA	CTTTCTCAG	ATTGATAGGC	ACATTCCCAT	3960
AGTCTACAAC	ATCAGCACAC	CAAGAAGAGG	TTGCCCCCTT	TGTGACTCTG	TAATCATGAG	4020
CTACGATGAG	AGCGATGGCC	TGCAGTCTCT	CATGTATGTC	ACTAATGAAA	GGGTGCAGAC	4080
CAACCTCTTT	TTAGATAAGT	CACCTTTCTT	TGATAATAAC	AACCTACACA	TTCATTATTT	4140
GTGGCTGAGG	GACAACGGGA	CCGTAGTGGA	GATAAGGGGC	ATGTATAGAA	GACGCGCAGC	4200
CAGTGCTTTG	TTTCTAATTC	TCTCTTTTAT	TGGGTTCTCG	GGGGTTATCT	ACTTTCTTTA	4260
CAGACTGTTT	TCCATCCTTT	ATTAGACGGT	CAATAAAGCG	TAGATTTTTA	AAAGGTTTCC	4320

180

006299: 6474999

TGTGCATTCT	TTTTGTATGG	GCATATACTT	GGCAAGAAAT	CCGAGCACCT	CAGAAAAGTGG	4380
ATTGCCGTCA	CATATCAGTT	CGACCACCCC	TGCACCTAGC	CATGCGGCGC	TTTGACGGTC	4440
TTTGGGGCTA	CACATCATAA	AGTACTTTTC	CATGGCTTCT	ATAAGCACCT	TGGAACAATC	4500
TGGGGGTTGG	CGAATGGGTT	CCCTAAACGG	GAAATCCTCT	ATGGTATTCA	GGCAGAAGAC	4560
CGCGTCCTCC	ACCCGACGTT	TGAGTCTTTC	TAGCAGAGCG	CCGAAGAACT	CCCGCTCGTG	4620
TGTTTTTCGA	GGGGCAAGTT	CTGCGCCGTA	CAGCGATGAG	AAACACGACA	CGATGTTTTTC	4680
CAGCCCCATG	CTGCGCAGCA	ACACGTGCTT	CAGGAACAGG	TGTTGTAGCC	GGTTCAGTTT	4740
TAGCTTGGGT	AGAAAAGTTA	TCGAGTTGTT	AGCACGCTCC	ATGATGGTAA	CGGTGTTGAA	4800
GTCACAGACC	GGGCTTTCTC	CGAGTCTCGG	CCGCCTGAGT	CCAATCATGT	AGAACATAGA	4860
CGCGGCCTCG	TTGTCTGTGT	TAAGTGACAC	GATATCCCGT	TCGCAAACCT	GTGCGATGTT	4920
GTGTTTCAGT	ATAGATCTGG	TCTGACCGGC	ACGGGGTGTT	ATGGGGTGAC	GCGGTAAAGG	4980
CGACTCTGGG	TCAAACACCT	TTATGCGGTT	GGCGGCCTCG	TCGATGACGA	CACGCTTGTT	5040
CGCGGCGTGT	ATGGGGACGC	GACGGCATCC	CGCTGGCAGA	TCTATAATCT	TAAAGTTGGT	5100
ATAAGACTGG	TCGCTCGTTA	TGGCCAGCCG	GCACTCCGGT	AGTATCTGCG	TGTCCTCGAA	5160
TTCGTGGCCG	CGTACGACTG	GCTTGAGTGT	CAGGTAAACG	CCAAGAGATG	CGGTCTCTTC	5220
GCCTACGCAC	AAGTGGCTTC	TTAACGCGTA	GGGGTCCGGT	GAGAGCATGA	TCCGTAGCAA	5280
CGATAGTTCC	GGGTGCCTAG	CCGCGTAGAG	TGGCAGGGTA	GACGAGTCCG	GAGTCCCAAA	5340
CTTTTCGAAC	AACAGTGGCA	TCGGGACTTC	AGGATTAGAG	ACTCCCACCA	TGGCCGCCAC	5400
CGCCGGAGAG	GTCAAGACGT	GAAACACGCG	CTCGCCTGTG	GACAGGCGCG	CCGCGCCCTC	5460
TACTAGACTA	GCCTTCACGT	CCGGAACTCG	TAACATAGCT	TAGACCAGCG	GACGGACGCA	5520
ACGTACGCGG	GGATCGGCTG	GCGGTGTCTG	CTCGTTGGAC	GCGGCCGTTC	GGTGGCGCCA	5580
GTGCAGGCCT	AGTTTGCGAA	TGGCGTGACG	GACAATTTGT	GGTTTTAGAG	CGGCGAACCG	5640
ATGACCCGTG	GTGGCGACGA	ACGAAATGAA	GTTTGCATTG	CGGCCCAACT	CGTCTAGCCT	5700
GGTCTTCTTG	TTTCGGGCAT	AGATTTTCGG	GATTAGGTTA	CACTTTTAT	ATCCAGTAC	5760
TGCGCACTCG	TGTTTGCTTT	TAGTGTGACT	GATTATCTTC	TTTGAGAAGT	CAAACAGGCC	5820
CCGGGCGGCG	GCTCGCCTAA	TGCAAGCCAC	GTCAAGCCTG	AGAAACGAAC	AGCATTCCAC	5880
CAGACACTCC	AGGAACCTTT	TGTGTAGCGT	CTGTATTTGG	GAACGGTTTC	TGTGCTCAAG	5940
TAGGGAGAAT	ATTCTATTTT	TGTTTCCGTC	GATGCGCGCG	TGCTGGTCCG	TGAGAATGGG	6000
CGCCAGCTCG	TGGCGAATCT	GTTCCACAAG	AGGCTGCCCC	TACACTTTAG	AAATCGTGGC	6060
TGTCGCGGCC	TTAAACCAGG	ACACGTTTAG	CCCATCCTTG	CTGGAGACCA	CAGATGGAAA	6120
GTTTGTGGTC	CAAAAATACGT	TTTTTCGCCC	CATTCTCACC	ATGTACTGGT	TTTCCAGTCC	6180
GTGCAGGTCC	AACGTGGAGT	TCCAATTTGC	TATCGATACA	GGAAATATGT	GCCTGATTGG	6240
CAGAAAGCAT	TTCAGCGTAC	CCATTGCGAA	GAGAAAGTGC	AGCATGTCCC	CACTGATGTT	6300
GATGTTTATT	GCGGTGCCTT	GACACATGTT	GTCGGAAAAA	AACACGCTTA	TGGTAAAAGA	6360

AGGTTCCCTTT	ACGGAGTACT	TTCTGATAAC	AAAATTGTTG	GTCAATCTGG	GGATGTTTAA	6420
AATAGTCTTT	TGCAGGGTGT	TAGGAACGTG	GCAGCTTATC	TTAGTGTTAA	TCACCATGTT	6480
GGTGTTGAAT	ATGGTGATCT	TGAAGTTTTT	CAAAGTACG	TGTTTTGTGG	GTTCCAGCAT	6540
GTCTGACACT	GTAGAGCTGC	CCAGAGTCCG	CGCTCCGTG	GCCGCGTATC	GTTGGAAGCA	6600
CGCCTGCAAA	TTTCCTTTCA	TGGCTGCTCG	CCGGTCTTTC	GGCGCGTACC	GGATTCTTGA	6660
AAGCGTCGCC	GCCAGGAGAC	GCGGTGTCTC	GTGGGTGCCT	AAAAAGTTTG	CGCAGGGGGT	6720
CAGTCCGCTG	CACGAGTGGC	CGATGCAGTC	TGCCACTGCC	ATACACATGA	CGAGTCTGTA	6780
GATGGCCGGT	GTGCCCCGAT	ACACTAGATA	GTAGGTACAA	TCTGGGGTAC	TGACGACCAC	6840
CCTGTATGGC	TTTGGTCCGG	GGTCTTTCG	TTGGATTTT	ACGTGCAGAC	GGGACACGAG	6900
CTGGTTTAGA	GCCAGCTGAA	AGCCCACCAG	ATCCCGTCCG	TTAACCTTGA	CGTCTGGTG	6960
CTTACTCTGT	TTGACAGGT	TCTTCAGCAC	GGTGGGCAGT	CGCTCTACGT	TGTGAGCGAT	7020
GGCACGGCGC	AGCGAGACCA	GCTCTCCGTG	CCACCCCCAC	GTGGCCATGA	AGCTGCTGAT	7080
GTTAAACTTT	AAAAAATGTA	GCTGTGCGTC	TGGGGATGCG	GGTGGCATT	TTGAAAACGA	7140
GAGATGCTTC	AGGCTCTCCA	GGAGTGCAAA	ATAATTTTGA	TAGATTGTGG	GTTGTAGACT	7200
ATGGGGCAAC	ACCGCCAGAA	ACGCATGAAA	ACACTGTTTC	AACTCCAGA	ACTCCAGGTA	7260
CCTGCACACT	ATCCTGAACA	TGGCTTTGTA	ACATATGGTG	CACGTTAGTA	GCGCGGGAAG	7320
ATACAGCGAG	CGTAGCTCCC	TGAATTCGCA	GGGTTTATCA	CAATCATCGG	TAAGTTCCCA	7380
TGATCCCACC	GCAGGTAGGT	AGTTGTGCGT	GTCTATCTGT	CCGCGCGTAA	ACACTCCACC	7440
ACCGTCAATT	ATTAAACCTT	CGCCGCTGTA	CCGTGACCC	ACTTTTCCCA	AAAGAGTCCC	7500
TTCTTGATGT	ATAAAAGGGT	GGAGGCGTTC	CCCCAGGAGT	AGTCTGCGTA	TCGCTCTGCA	7560
GGCGAAAAAG	GTGGGCTCGG	GCTGCATCAT	CTTATCAAGA	CTTCTAAGG	TCAGCTCTGC	7620
CTGCAGGTGC	GAGTTGGTGG	CCAGACAGCA	GAATATTTCC	AGCTGTGATT	CCCAAGTCGC	7680
TTGATAACAC	GTGGTCTGCG	GACTCGTCGT	CAGGGAGGCG	CTCGTGGCA	GTAGTAGGGG	7740
GCCCTCGAGC	GCTGCCATGG	AGGCGACCTT	GGAGCAACGA	CCTTTCCCGT	ACCTCGCCAC	7800
GGAGGCCAAC	CTCCTAACGC	AGATTAAGGA	GTGGGCTGCC	GACGGACTCT	TCAAGAGCTT	7860
TCAGCTATTG	CTCGGCAAGG	ACGCCAGAGA	AGGCAGTGTC	CGTTTCGAAG	CGCTACTGGG	7920
CGTATATACC	AATGTGGTGG	AGTTTGTTAA	GTTTCTGGAG	ACCGCCCTCG	CCGCCGCTTG	7980
CGTCAATACC	GAGTTCAAGG	ACCTGCGGAG	AATGATAGAT	GGAAAAATAC	AGTTTAAAAAT	8040
TTCAATGCCC	ACTATTGCCC	ACGGAGACGG	GAGGAGGCCC	AACAAGCAGA	GACAGTATAT	8100
CGTCATGAAG	GCTTGCAATA	AGCACCACAT	CGGTGCGGAG	ATTGAGCTTG	CGGCCGCAGA	8160
CATCGAGCTT	CTCTTCGCCG	AGAAAGAGAC	GCCCTTGGAC	TTACACAGAGT	ACGCGGGTGC	8220
CATCAAGACG	ATTACGTCGG	CTTTGCAGTT	TGGTATGGAC	GCCCTAGAAC	GGGGGCTAGT	8280
GGACACGGTT	CTCGCAGTTA	AACTTCGGCA	CGCTCCACCC	GTCTTTATTT	TAAAGACGCT	8340
GGGCGATCCC	GTCTACTCTG	AGAGGGGCCCT	CAAAAAGGCC	GTCAAGTCTG	ACATGGTATC	8400

006290" 6 2 F 2 0 9 6 0

CATGTTCAAG	GCACACCTCA	TAGAACATTC	ATTTTTTCTA	GATAAGGCCG	AGCTCATGAC	8460
AAGGGGGAAG	CAGTATGTCC	TAACCATGCT	CTCCGACATG	CTGGCCGCGG	TGTGCGAGGA	8520
TACCGTCTTT	AAGGGTGTCA	GCACGTACAC	CACGGCCTCT	GGGCAGCAGG	TGGCCGGCGT	8580
CCTGGAGACG	ACGGACAGCG	TCATGAGACG	GCTGATGAAC	CTGCTGGGGC	AAGTGGAAA	8640
TGCCATGTCC	GGGCCCCGCG	CTACGCCAG	CTACGTTGTC	AGGGGTGCCA	ACCTCGTCAC	8700
CGCCGTTAGC	TACGGAAGGG	CGATGAGAAA	CTTTGAACAG	TTTATGGCAC	GCATAGTGGA	8760
CCATCCCAAC	GCTCTGCCGT	CTGTGGAAGG	TGACAAGGCC	GCTCTGGCGG	ACGGACACGA	8820
CGAGATTCAG	AGAACCCGCA	TCGCCGCTC	TCTCGTCAAG	ATAGGGGATA	AGTTTGTGGC	8880
CATTGAAAGT	TTGCAGCGCA	TGTACAACGA	GACTCAGTTT	CCCTGCCCAC	TGAACCGGCG	8940
CATCCAGTAC	ACCTATTTCT	TCCCTGTTGG	CCTTCACCTT	CCCGTGCCCC	GCTACTCGAC	9000
ATCCGTCTCA	GTCAGGGGCG	TAGAATCCCC	GGCCATCCAG	TCGACCGAGA	CGTGGGTGGT	9060
TAATAAAAC	AACGTGCCTC	TTTGCTTCGG	TTACCAAAAC	GCCCTCAAAA	GCATATGCCA	9120
CCCTCGAATG	CACAACCCCA	CCCAGTCAGC	CCAGGCACTA	AACCAAGCTT	TTCCCGATCC	9180
CGACGGGGGA	CATGGGTACG	GTCTCAGGTA	TGAGCAGACG	CCAAACATGA	ACCTATTCAG	9240
AACGTTCCAC	CAGTATTACA	TGGGGAAAAA	CGTGGCATT	GTTCCCGATG	TGGCCCAAAA	9300
AGCGCTCGTA	ACCACGGAGG	ATCTACTGCA	CCCAAGCTCT	CACCGTCTCC	TCAGATTGGA	9360
GGTCCACCCC	TTCTTTGATT	TTTTTGTGCA	CCCCTGTCT	GGAGCGAGAG	GATCGTACCG	9420
CGCCACCCAC	AGAACAATGG	TTGGAAATAT	ACCACAACCG	CTCGCTCCAA	GGGAGTTTCA	9480
GGAAAGTAGA	GGGGCGCAGT	TCGACGCTGT	GACGAATATG	ACACACGTCA	TAGACCAGCT	9540
AACTATTGAC	GTCATACAGG	AGACGGCATT	TGACCCCGCG	TATCCCCTGT	TCTGCTATGT	9600
AATCGAAGCA	ATGATTCACG	GACAGGAAGA	AAAATTCGTG	ATGAACATGC	CCCTCATTGC	9660
CCTGGTCATT	CAAACCTACT	GGGTCAACTC	GGGAAACTG	GCGTTTGTGA	ACAGTTATCA	9720
CATGGTTAGA	TTCATCTGTA	CGCATATTGG	GAATGGAAGC	ATCCCTAAGG	AGGCGCACGG	9780
CCACTACCGG	AAAATCTTAG	GCGAGCTCAT	CGCCCTTGAG	CAGGCGCTTC	TCAAGCTCGC	9840
GGGACACGAG	ACGGTGGGTC	GGACGCCGAT	CACACATCTG	GTTTCGGCTC	TCCTCGACCC	9900
GCATCTGCTG	CCTCCCTTTG	CCTACCACGA	TGCTTTTACG	GATCTTATGC	AGAAGTCATC	9960
CAGACAACCC	ATAATCAAGA	TCGGGGATCA	AAACTACGAC	AACCCTCAA	ATAGGGCGAC	10020
ATTCATCAAC	CTCAGGGGTC	GCATGGAGGA	CCTAGTCAAT	AACCTTGTTA	ACATTTACCA	10080
GACAAGGGTC	AATGAGGACC	ATGACGAGAG	ACACGTCCTG	GACGTGGCGC	COCTGGACGA	10140
GAATGACTAC	AACCCGGTCC	TCGAGAAGCT	ATTCTACTAT	GTTTTAATGC	CGGTGTGCAG	10200
TAACGGCCAC	ATGTGCGGTA	TGGGGGTGCA	CTATCAAAAC	GTGGCCCTGA	CGCTGACTTA	10260
CAACGGCCCC	GTCTTTGCGG	ACGTCGTGAA	CGCACAGGAT	GATATTCTAC	TGCACCTGGA	10320
GAACGGAACC	TTGAAGGACA	TTCTGCAGGC	AGGCGACATA	CGCCCGACGG	TGGACATGAT	10380
CAGGGTGCTG	TGCACCTCGT	TTCTGACGTG	CCCTTTCGTC	ACCCAGGCCG	CTCGCGTGAT	10440

CACAAAGCGG	GACCCGGGCC	AGAGTTTTGC	CACGCACGAA	TACGGGAAGG	ATGTGGCGCA	10500
GACCGTGCTT	GTTAATGGGT	TTGGTGCGTT	CGCGGTGGCG	GACCGCTCTC	GCGAGGCGGC	10560
GGAGACTATG	TTTTATCCGG	TACCCTTTAA	CAAGCTCTAC	GCTGACCCGT	TGGTGGCTGC	10620
CACACTGCAT	CCGCTCCTGC	CAAACTATGT	CACCAGGCTC	CCCAACCAGA	GAAACGCGGT	10680
GGTCTTTAAC	GTGCCATCCA	ATCTCATGGC	AGAATATGAG	GAATGGCACA	AGTCGCCCCGT	10740
CGCGGGCGTAT	GCCGCGTCTT	GTCAGGCCAC	CCCGGGCGCC	ATTAGCGCCA	TGGTGAGCAT	10800
GCACCAAAAA	CTATCTGCCC	CCAGTTTCAT	TTGCCAGGCA	AAACACCGCA	TGCACCCTGG	10860
TTTTTGCCATG	ACAGTCGTCA	GGACGGACGA	GGTTCTAGCA	GAGCACATCC	TATACTGCTC	10920
CAGGGCGTCG	ACATCCATGT	TTGTGGGCTT	GCCTTCGGTG	GTACGGCGCG	AGGTACGTTT	10980
GGACGCGGTG	ACTTTTGAAA	TTACCCACGA	GATCGCTTCC	CTGCACACCG	CACTTGGCTA	11040
CTCATCAGTC	ATCGCCCCGG	CCCACGTGGC	CGCCATAACT	ACAGACATGG	GAGTACATTG	11100
TCAGGACCTC	TTTATGATTT	TCCCAGGGGA	CGCGTATCAG	GACCGCCAGC	TGCATGACTA	11160
TATCAAAATG	AAAGCGGGCG	TGCAAACCGG	CTCACC GGGA	AACAGAATGG	ATCACGTGGG	11220
ATACACTGCT	GGGGTTCCTC	GCTGCGAGAA	CCTGCCCGGT	TTGAGTCATG	GTCAGCTGGC	11280
AACCTGCGAG	ATAATTCCCA	CGCCGGTCAC	ATCTGACGTT	GCCTATTTCC	AGACCCCCAG	11340
CAACCCCCGG	GGGCGTGCGG	CGTCGGTCGT	GTCGTGTGAT	GCTTACAGTA	ACGAAAGCGC	11400
AGAGCGTTTTG	TTCTACGACC	ATTCAATACC	AGACCCCGCG	TACGAATGCC	GGTCCACCAA	11460
CAACCCGTGG	GCTTCGCAGC	GTGGCTCCCT	CGGCGACGTG	CTATACAATA	TCACCTTTTCG	11520
CCAGACTGCG	CTGCCGGGCA	TGTACAGTCC	TTGTGGGCAG	TTCTTCCACA	AGGAAGACAT	11580
TATGCGGTAC	AATAGGGGGT	TGTACACTTT	GGTTAATGAG	TATTCTGCCA	GGCTTGCTGG	11640
GGCCCCCGCC	ACCAGCACTA	CAGACCTCCA	GTACGTCGTG	GTCAACGGTA	CAGACGTGTT	11700
TTTGGACCAG	CCTTGCCATA	TGCTGCAGGA	GGCCTATCCC	ACGCTCGCCG	CCAGCCACAG	11760
AGTTATGCTT	GCCGAGTACA	TGTCAAACAA	GCAGACACAC	GCCCCAGTAC	ACATGGGGCCA	11820
GTATCTCATT	GAAGAGGTGG	CGCCGATGAA	GAGACTATTA	AAGCTCGGAA	ACAAGGTGGT	11880
GTATTAGCTA	ACCCTTCTAG	CGTTGGCTAG	TCATGGCACT	CGACAAGAGT	ATAGTGGTTA	11940
ACTTCACCTC	CAGACTCTTC	GCTGATGAAC	TGGCCGCCCT	TCAGTCAAAA	ATAGGGAGCG	12000
TACTGCCGCT	CGGAGATTGC	CACCGTTTAC	AAAATATACA	GGCATTGGGC	CTGGGGTGCG	12060
TATGCTCACG	TGAGACATCT	CCGGACTACA	TCCAAATTAT	GCAGTATCTA	TCCAAGTGCA	12120
CACTCGCTGT	CCTGGAGGAG	GTTGCGCCCG	ACAGCCTGCG	CCTAACGCGG	ATGGATCCCT	12180
CTGACAACCT	TCAGATAAAA	AACGTATATG	CCCCCTTTTT	TCAGTGGGAC	AGCAACACCC	12240
AGCTAGCAGT	GCTACCCCCA	TTTTTTAGCC	GAAAGGATTG	CACCATTGTG	CTCGAATCCA	12300
ACGGATTGGA	CCCCGTGTTC	CCCATGGTCG	TGCCGCAGCA	ACTGGGGCAC	GCTATTCTGC	12360
AGCAGCTGTT	GGTGTACCAC	ATCTACTCCA	AAATATCGGC	CGGGGCCCCG	GATGATGTAA	12420
ATATGGCGGA	ACTTGATCTA	TATACCACCA	ATGTGTCATT	TATGGGGCGC	ACATATCGTC	12480

TGGACGTTAGA	CAACACGGAT	CCACGTACTG	CCCTGCGAGT	GCTTGACGAT	CTGTCCATGT	12540
ACCTTTGTAT	CCTATCAGCC	TTGGTCCCCA	GGGGGTGTCT	CCGTCTGCTC	ACGGCGCTCG	12600
TGCGGCACGA	CAGGCATCCT	CTGACAGAGG	TGTTTGAGGG	GGTGGTGCCA	GATGAGGTGA	12660
CCAGGATAGA	TCTCGACCAG	TTGAGCGTCC	CAGATGACAT	CACCAGGATG	CGCGTCATGT	12720
TCTCCTATCT	TCAGAGTCTC	AGTTCTATAT	TTAATCTTGG	CCCCAGACTG	CACGTGTATG	12780
CCTACTCGGC	AGAGACTTTG	GCGGCCTCCT	GTTGGTATTC	CCCACGCTAA	CGATTTGAAG	12840
CGGGGGGGGT	ATGGCGTCAT	CTGATATTCT	GTCGGTTGCA	AGGACGGATG	ACGGCTCCGT	12900
CTGTGAAGTC	TCCCTGCGTG	GAGGTAGGAA	AAAACTACC	GTCTACCTGC	CGGACACTGA	12960
ACCCTGGGTG	GTAGAGACCG	ACGCCATCAA	AGACGCCTTC	CTCAGCGACG	GGATCGTGGA	13020
TATGGCTCGA	AAGCTTCATC	GTGGTGCCCT	GCCCTCAAT	TCTCACAACG	GCTTGAGGAT	13080
GGTGCTTTTT	TGTTATTGTT	ACTTGCAAAA	TTGTGTGTAC	CTAGCCCTGT	TTCTGTGCCC	13140
CCTTAATCCT	TACTTGGTAA	CTCCCTCAAG	CATTGAGTTT	GCCGAGCCCG	TTGTGGCACC	13200
TGAGGTGCTC	TTCCACACCC	CGGCTGAGAT	GTCTCGCGGT	TGCGATGACG	CGATTTTCTG	13260
TAAACTGCCC	TATACCGTGC	CTATAATCAA	CACCACGTTT	GGACGCATTT	ACCCGAACTC	13320
TACACGCGAG	CCGGACGGCA	GGCCTACGGA	TTACTCCATG	GCCCTTAGAA	GGGCTTTTGC	13380
AGTTATGGTT	AACACGTCAT	GTGCAGGAGT	GACATTGTGC	CGCGGAGAAA	CTCAGACCGC	13440
ATCCCCTAAC	CACACTGAGT	GGGAAAATCT	GCTGGCTATG	TTTTCTGTGA	TTATCTATGC	13500
CTTAGATCAC	AACTGTCACC	CGGAAGCACT	GTCTATCGCG	AGCGGCATCT	TTGACGAGCG	13560
TGACTATGGA	TTATTATCTT	CTCAGCCCCG	GAGCGTGCCC	TCGCTACCC	CTTGCGACGT	13620
GTCGTGGGAA	GATATCTACA	ACGGGACTTA	CCTAGCTCGG	CCTGGAAACT	GTGACCCCTG	13680
GCCCAATCTA	TCCACCCCTC	CCTTGATTCT	AAATTTTAAA	TAAAGGTGTG	TCACTGGTTA	13740
CACCACGATT	AAAAACCACT	CACTGAGATG	TCTTTTAAAC	CGCTAAGGGA	TTATACCGGG	13800
ATTTAAAACC	GCCCACTGAT	TTTTTTACGC	TAAGAGTTGG	GTGCTTGGGG	GGTTTTGCAT	13860
TGCTCTGTTG	TAAACTATAT	ATAAGTTAAA	CCAAAATTCG	CAGGGAGACA	AGGTGACGGT	13920
GGTGAGAACT	CAGTTGAGAG	TCAGAGAATA	CAGTGCTAAT	CAGGGTAGAT	GAGCATGACT	13980
TTCCCCGTCT	CCAGTCACCG	GAGGAATGGT	GGACGGCTCC	GTCCTGGTGC	GAATGGCCAC	14040
CAAGCCTCCC	GTGATTGGTC	TTATAACAGT	GCTCTTCCTC	CTAGTCATAG	GCGCCTGCGT	14100
CTACTGCTGC	ATTGCGGTGT	TCCTGGCGGC	TCGACTGTGG	CGCGCCACCC	CACTAGGCAG	14160
GGCCACCGTG	GCGTATCAGG	TCCTTCGCAC	CCTGGGACCG	CAGGCCGGGT	CACATGCACC	14220
GCCGACGGTG	GGCATAGCTA	CCCAGGAGCC	CTACCGTACA	ATATACATGC	CAGATTAGAA	14280
CGGGGTGTGT	GCTATAATGG	ATGGCTATGG	GGGGGGGCTG	TAGATAATTG	AGCGCTGTGC	14340
TTTTATTGTG	GGGATATGGG	CTTGATACATG	TGTCTATCAT	CGGTAGCCAT	AAAATGGGCC	14400
ATGACAACCTG	CCACAAGTAA	GTCGTCCGAC	ATGTGCTTTT	GCTTGGCGCT	GTATGACTGC	14460
CCTCCATCCC	TAAGCGGGAC	GCACTTGATC	GCGCGGACCT	GTTCTACCAG	GTAGGTCACC	14520

GGGTCAAATG	ATATTTTGAT	GCTGTTGGAC	ACCACCGTCT	GGCTGGCGCT	CAGGGTGCCG	14580
GAGTTCAGAG	CGTAGATGAA	TGCTCTCAAAC	GCGGAGGATT	TCTCGCCTCC	CAACATGTAA	14640
ATTGGCCACT	GCAGGGCGCT	GCTCTTGTC	GTATAGTGTA	GAAAATGTAT	GGGGAGCGGG	14700
CATATTTCTG	TAAGGACGGT	TGCAATGGCC	ACCCAGAAAT	CTTGGCTGCT	GTTGCCTTCG	14760
ACCGCCGCGT	TCACGCGCTC	AATTGTGGTG	TGGAGCACAG	CGATCGCCTT	AATCATCGTG	14820
CATGCGCAGG	ACGCTATCTC	GTAAGCAGCT	GCGCCAGTGA	GGTCGCGCAG	GAAGAAATGC	14880
TCCATGCCCA	ATATGAGGCT	TCTGGTGGGA	GTCTGAGTAC	TCGTGACAAC	GGCGCCACG	14940
CCAGTACCGG	ACGCCTCCGT	GTTGTTTCGTA	TACGCGGGGT	CGATGTAAAC	AAACAGCTGT	15000
TTTCCAAGGC	ACTTCTGAAC	CTCCTGGGCG	GTGGTGTCTA	CCCGACACAT	GTCAAACCTGT	15060
GTCAGCGCTG	CGTCACCCAC	CACGCGGTAA	AGCGTAGCAT	TTGACGACGC	TGCTCCCTCG	15120
CCCATTAGTT	CGGTGTCGAA	TGCCCCCTCC	ATAAGCAGGT	TGGTGGTGGT	TTTGATGGAT	15180
TCGTGATGG	TGATGTACGT	CGGAATGTGC	AGTCTGTAC	AAGGACAGGA	CACTAGTGCG	15240
TCTTGACAGG	GGAAATCTTC	TCGGTGGTCC	GCACACACGT	AACTGACCAC	ATTCAGCATC	15300
TTTTCTGGG	CGTTCCTGAG	GTTAAGCAGG	AACTCGTGG	AGCGGTCTGA	CGAGTTCACG	15360
GATGATATAA	ATATAAGCTT	GGCGTCTTTC	TGAAGCATGA	AACCCAGAAT	AGCCGGCAGT	15420
GCATCCTTTT	TAATAAAATT	CGCCTCGTCT	ACGTAGAGCA	GGTTAAAGGT	CTGTCCCCGA	15480
ATGCTCTGCA	GACACGGAAA	GACACAAAAG	AGGGGCTCAT	AAGCGGCTAA	CAGTAAAGGA	15540
GAGGAGGCGA	ACAGTGCGTG	GCTCTTGGTT	CTTGGGAATA	AAAGGGGGCG	TGTGTGCCGA	15600
TCGATCGTAT	GGGTGAGCCA	GTGGATCCTG	GACATGTGGT	GAATGAGAAA	GATTTTGAGG	15660
AGTGTGAACA	ATTTTTCAGT	CAACCCCTTA	GGGAGCAAGT	GGTCGCGGGG	GTCAGGGCAC	15720
TCGACGGCCT	CGGTCTCGCT	GACTCTCTAT	GTCACAAAAC	AGAAAGACTC	TGCCTGCTGA	15780
TGGACCTGGT	GGGCACGGAG	TGCTTTGCGA	GGGTGTGCCG	CCTAGACACC	GGTGCGAAAT	15840
GAAGAGTGTG	GCGAGTCCCT	TATGTCAGTT	CCACGGCGTG	TTTTGCCTGT	ACCAGTGTCG	15900
CCAGTGCCCTG	GCATACCACG	TGTGTGATGG	GGGCGCCGAA	TGCGTTCTCC	TGCATACGCC	15960
GGAGAGCGTC	ATCTGCGAAC	TAACGGGTAA	CTGCATGCTC	GGCAACATTC	AAGAGGGCCA	16020
GTTTTTTAGGG	CCGGTACCGT	ATCGGACTTT	GGATAACCAG	GTTGACAGGG	ACGCATATCA	16080
CGGGATGCTA	GCGTGTCTGA	AACGGGACAT	TGTGCGGTAT	TTGCAGACAT	GGCCGGACAC	16140
CACCGTAATC	GTGCAGGAAA	TAGCCCTGGG	GGACGGCGTC	ACCGACACCA	TCTCGGCTAT	16200
TATAGATGAA	ACATTCGGTG	AGTGTCTTCC	CGTACTGGGG	GAGGCCCAAG	GCGGGTACCC	16260
CCTGGTCTGT	AGCATGTATC	TGCACGTTAT	CGTCTCCATC	TATTCGACAA	AAACGGTGTA	16320
CAACAGTATG	CTATTTAAAT	GCACAAAGAA	TAAAAAGTAC	GACTGCATTG	CCAAGCGGGT	16380
GCGGACAAAA	TGGATGCGCA	TGCTATCAAC	GAAAGATACG	TAGGTCCTCG	CTGCCACCGT	16440
TTGGCCACAG	TGGTGCTGCC	TAGGACCTTT	CTGCTGCATC	ACGCCATACC	CCTGGAGCCC	16500
GAGATCATCT	TTTCCACCTA	CACCCGGTTC	AGCCGGTCGC	CAGGGTCATC	CCGCCGGTTG	16560

GTGGTGTGTG GGAAACGTGT CCTGCCAGGG GAGGAAAACC AACTTGCGTC TTCACCTTCT	16620
GGTTTGGCGC TTAGCTGCC TCTGTTTCC CACGATGGGA ACTTTCATCC ATTTGACATC	16680
TGGTACTGC GCATTTCTG CCCTGGTTCT AATCTTAGTC TTACTGTCAG ATTTCTCTAT	16740
CTATCTCTGG TGGTGGCTAT GGGGGCGGGA CGGAATAATG CGCGGAGTCC GACCGTTGAC	16800
GGGGTATCGC CGCCAGAGGG CGCCGTAGCC CACCCTTTGG AGGAACTGCA GAGGCTGGCG	16860
CGTGCTACGC CGGACCCGGC ACTCACCCTT GGACCCTGAC AGGTCCTGAC CGGCCTTCTC	16920
CGCGCAGGGT CAGACGGAGA CCGCCTTACT CACCACATGG CGCTCGAGGC TCCGGGAACC	16980
GTGCGTGGAG AAAGCCTAGA CCCGCTTGT TCACAGAAGG GGCCAGCGCG CACACGCCAC	17040
AGGCCACCCC CCGTGCGACT GAGCTTCAAC CCCGTCAATG CCGATGTACC CGCTACCTGG	17100
CGAGACGCCA CTAACGTGTA CTCGGGTGCT CCCTACTATG TGTGTGTTTA CGAACGCGGT	17160
GGCCGTCAGG AAGACGACTG GCTGCCGATA CCACTGAGCT TCCAGAAGA GCCCGTGCCC	17220
CCGCCACCGG GCTTAGTGTT CATGGACGAC TTGTTTATTA ACACGAAGCA GTGCGACTTT	17280
GTGGACACGC TAGAGGCCGC CTGTCGCACG CAAGGCTACA CGTTGAGACA GCGCGTGCCT	17340
GTCGCCATTC CTCGCGACGC GGAAATCGCA GACGAGTTA AATCGCACTT TTTAGAGGCG	17400
TGCCTAGTGT TACGGGGGCT GGCTTCGGAG GCTAGTCCCT GGATAAGAGC TGCCACGTCC	17460
CCGCCCTTG GCCGCCACGC CTGCTGGATG GACGTGTTAG GATTATGGGA AAGCCGCCCC	17520
CACACTCTAG GTTTGGAGTT ACGCGGCGTA AACTGTGGCG GCACGGACGG TGA CTGGTTA	17580
GAGATTTTAA AACAGCCCCG TGTGCAAAAG ACAGTCAGCG GGAGTCTTGT GGCATGCGTG	17640
ATCGTCACAC CCGCATTGGA AGCCTGGCTT GTGTTACCTG GGGGTTTTGC TATTAAAGCC	17700
CGCTATAGGG CGTCGAAGGA GGATCTGGTG TTCATTCGAG GCGCTATGG CTAGCCGGAG	17760
GCGCAAATT CGGAATTTCC TAAACAAGGA ATGCATATGG ACTGTTAACC CAATGTCAGG	17820
GGACCATATC AAGGTCTTTA ACGCCTGCAC CTCTATCTCG CCGGTGTATG ACCCTGAGCT	17880
GGTAACCAGC TACGCACTGA GCGTGCCTGC TTACAATGTG TCTGTGGGTA TCTTGCTGCA	17940
TAAAGTCATG GGACCGTGTG TGGCTGTGGG AATTAACGGA GAAATGATCA TGTACGTCGT	18000
AAGCCAGTGT GTTTCTGTGC GGCCCGTCCC GGGGCGCGAT GGTATGGCGC TCATCTACTT	18060
TGGACAGTTT CTGGAGGAAG CATCCGGA CTGAGTTTCCC TACATTGCTC CCGCGCGTC	18120
GCGCGAACAC GTACCTGACC TGACCAGACA AGAATTAGTT CATACCTCCC AGGTGGTGCG	18180
CCGCGGCGAC CTGACCAATT GCACTATGGG TCTCGAATTC AGGAATGTGA ACCCTTTTGT	18240
TTGGCTCGGG GCGGATCGG TGTGGCTGCT GTTCTTGGGC GTGGACTACA TGGCGTCTG	18300
TCCGGGTGTC GACGGAATGC CGTCGTTGGC AAGAGTGGCC GCCCTGCTTA CCAGGTGCGA	18360
CCACCCAGAC TGTGTCCACT GCCATGGACT CCGTGGACAC GTTAATGTAT TTCGTGGGTA	18420
CTGTTCTGCG CAGTCGCCGG GTCTATCTAA CATCTGTCCC TGTATCAAAT CATGTGGGAC	18480
CGGGAATGGA GTGACTAGGG TCACTGGAAA CAGAAATTTT CTGGGTCTTC TGTTCGATCC	18540
CATTGTCCAG AGCAGGGTAA CAGCTCTGAA GATAACTAGC CACCCAACCC CCACGCACGT	18600

CGAGAATGTG CTAACAGGAG TGCTCGACGA CGGCACCTTG GTGCCGTCCG TCCAAGGCAC 18660
CCTGGGTCCT CTTACGAATG TCTGACTACT TCAGCCGCTT GCTGATATAT GAGTGTAATA 18720
AACTTAAGGC CCTGGGCTTA CGTTCTTATT GAAGCATGTT GCGCACATCA GCGAGCTGGA 18780
CCGTCCTCCG GGTGCGGTGT AGATTATGGT TCCGTTCTCC TTCTTGATGT TTAAATTTTT 18840
GGGGGGGAAC CACCGACAAA GCGTCTTTAT GATTTCCGCG AACACGGAGT TGGCTACGTG 18900
CTTTTGTTGG GCTACGTACC CAATGTTAAT GTTCTCTACG GATGCCAGTA GCATGCTGAT 18960
GATCGCCACC ACTATCCATG TCTTTCCGTG TCTCCTTGGT ATTAGGAATA CGCTTGCCCTT 19020
TTGCTTAAAC GTCTGTAAAA CACTGTTTGG AGTTTCAAAT AAACCGAAGT ACTGCTTAAA 19080
CAATCCAAAC AACTGGTGCG TCTTTTGTGG GGCCTTGATT GAAACCAAAA AGAAAAAAGT 19140
GTGCATTACT AGCTGCTGTT GGAAGGGCTC CAGCCAGTGC ACCCCGGGAA CGTAACAGCC 19200
GTTCAGAAAG GACGAAAGGT TAACCAGAAA AGCCTGAAGT TCGCGGTAGA CAGAGCAGGC 19260
GTGCAGGGAG TCGTGTGTTT TTCTGCCCCG CTGGTACTCG ACCAGTTGAT CGGCCGTGGA 19320
GACGTGCGCG TCCTCGCGCA CACACCGCAT CTGCAAGTAT GTTGATAGGG ACTCCAATAG 19380
GCGCGGCTTT GCGGGGACGT TGTCTCGGA CGGTGTGGGG GTTCCCACGT CGGGATTGTC 19440
TGACGTGGGC GTGGCGGGAT GGTGCCGTGT GCAGTATGTT TCCAGGACCG AACTGTATGA 19500
GTTTATTCTG TGCACCACGC CAATAAAAGG GTGCGCCATC CGTGCCGTTT TGGGACAGTG 19560
TCGCGTGAAT GTCGGGGCAC TCAGTTCCCA CCTCTCTCCG GCGTCTTTGG CGGTCTCCTC 19620
CAGGTGCGCG GCAAGGCGCT CCCTGTGACG GCTGAGCAGC ATGTTTGCTT TGAGCTCGCT 19680
CGTGTCCGAG GGTGACCCGG AGGTGACCAG TAGGTACGTC AAGGGCGTAC AACTTGCCCT 19740
GGACCTTAGC GAGAACACAC CTGGACAATT TAAGTTGATA GAAACTCCCC TGAACAGCTT 19800
CCTCTTGTTT TCCAACGTGA TGCCCGAGGT CCAGCCAATC TGCAAGTGCC GGCCGGCCCTT 19860
GCGGCCAGAC TTTAGTAATC TCCACTTGCC TAGACTGGAG AAGCTTCAGA GAGTCCTCGG 19920
GCAGGGTTTC GGGGCGGCGG GTGAGGAAAT CGCACTGGAC CCGTCTGACG TAGAAACACA 19980
CGAAAAGGGC CAGGTGTTCT ACAACCACTA TGCTACCGAG GAGTGGACGT GGGCTTTGAC 20040
TCTGAATAAG GATGCGCTCC TTCGGGAGGC TGTAGATGGC CTGTGTGACC CCGGAACCTG 20100
GAAGGGTCTT CTTCCTGACG ACCCCCTTCC GTTGCTATGG CTGCTGTTCA ACGGACCCGC 20160
CTCTTTTTGT CGGGCCGACT GTTGCCGTGA CAAGCAGCAC TGCGGTTACC CCGGCCCGGT 20220
GCTACTTCCA GGTACATGT ACGCTCCCAA ACGGGATCTT TTGTCGTTCC TTAATCATGC 20280
CCTGAAGTAC ACCAAGTTTC TATACGGAGA TTTTCCGGG ACATGGGCGG CGGCTTGCCG 20340
CCCGCCATTC GCTACTTCTC GGATACAAAG GGTAGTGAGT CAGATGAAAA TCATAGATGC 20400
TTCCGACACT TACATTTCCC ACACCTGCCT CTTGTGTCAC ATATATCAGC AAAATAGCAT 20460
AATTGCGGGT CAGGGGACCC ACGTGGGTGG AATCCTACTG TTGAGTGGAA AAGGGACCCA 20520
GTATATAACA GGCAATGTTT AGACCCAAAG GTGTCCAAC ACAGGCGGACT ATCTAATCAT 20580
CCCATCGTAT GACATACCGG CGATCATCAC CATGATCAAG GAGAATGGAC TCAACCAACT 20640

006290"6440960

188

CTAAAAGAGA GTTTATTAAG TCGGCTCTGG AGGCCAACAT CAACAGGAGG GCAGCTGTAT 20700
CGCTATTTGA 20710

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4131 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..4131
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATG GAG GCG ACC TTG GAG CAA CGA CCT TTC CCG TAC CTC GCC ACG GAG 48
Met Glu Ala Thr Leu Glu Gln Arg Pro Phe Pro Tyr Leu Ala Thr Glu
1 5 10 15

GCC AAC CTC CTA ACG CAG ATT AAG GAG TCG GCT GCC GAC GGA CTC TTC 96
Ala Asn Leu Leu Thr Gln Ile Lys Glu Ser Ala Ala Asp Gly Leu Phe
20 25 30

AAG AGC TTT CAG CTA TTG CTC GGC AAG GAC GCC AGA GAA GGC AGT GTC 144
Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Ser Val
35 40 45

CGT TTC GAA GCG CTA CTG GGC GTA TAT ACC AAT GTG GTG GAG TTT GTT 192
Arg Phe Glu Ala Leu Leu Gly Val Tyr Thr Asn Val Val Glu Phe Val
50 55 60

AAG TTT CTG GAG ACC GCC CTC GCC GCC GCT TGC GTC AAT ACC GAG TTC 240
Lys Phe Leu Glu Thr Ala Leu Ala Ala Ala Cys Val Asn Thr Glu Phe
65 70 75 80

AAG GAC CTG CGG AGA ATG ATA GAT GGA AAA ATA CAG TTT AAA ATT TCA 288
Lys Asp Leu Arg Arg Met Ile Asp Gly Lys Ile Gln Phe Lys Ile Ser
85 90 95

ATG CCC ACT ATT GCC CAC GGA GAC GGG AGG AGG CCC AAC AAG CAG AGA 336
Met Pro Thr Ile Ala His Gly Asp Gly Arg Arg Pro Asn Lys Gln Arg
100 105 110

CAG TAT ATC GTC ATG AAG GCT TGC AAT AAG CAC CAC ATC GGT GCG GAG 384
Gln Tyr Ile Val Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125

ATT GAG CTT GCG GCC GCA GAC ATC GAG CTT CTC TTC GCC GAG AAA GAG 432
Ile Glu Leu Ala Ala Ala Asp Ile Glu Leu Leu Phe Ala Glu Lys Glu
130 135 140

ACG CCC TTG GAC TTC ACA GAG TAC GCG GGT GCC ATC AAG ACG ATT ACG 480
Thr Pro Leu Asp Phe Thr Glu Tyr Ala Gly Ala Ile Lys Thr Ile Thr
145 150 155 160

006290-04F20950

THE

TGG Ser	GCT Ala	TTG Leu	CAG Gln	TTT Phe 165	GGT Gly	ATG Met	GAC Asp	GCC Ala	CTA Leu 170	GAA Glu	CGG Arg	GGG Gly	CTA Leu	GTG Val 175	GAC Asp	528
ACG Thr	GTT Val	CTC Leu	GCA Ala 180	GTT Val	AAA Lys	CTT Leu	CGG Arg	CAC His 185	GCT Ala	CCA Pro	CCC Pro	GTC Val 190	TTT Phe	ATT Ile	TTA Leu	576
AAG Lys	ACG Thr	CTG Leu 195	GGC Gly	GAT Asp	CCC Pro	GTC Val	TAC Tyr 200	TCT Ser	GAG Glu	AGG Arg	GGC Gly	CTC Leu 205	AAA Lys	AAG Lys	GCC Ala	624
GTC Val	AAG Lys 210	TCT Ser	GAC Asp	ATG Met	GTA Val	TCC Ser 215	ATG Met	TTC Phe	AAG Lys	GCA Ala	CAC His 220	CTC Leu	ATA Ile	GAA Glu	CAT His	672
TCA Ser 225	TTT Phe	TTT Phe	CTA Leu	GAT Asp	AAG Lys 230	GCC Ala	GAG Glu	CTC Leu	ATG Met	ACA Thr 235	AGG Arg	GGG Gly	AAG Lys	CAG Gln	TAT Tyr 240	720
GTC Val	CTA Leu	ACC Thr	ATG Met	CTC Leu 245	TCC Ser	GAC Asp	ATG Met	CTG Leu	GCC Ala 250	GCG Ala	GTG Val	TGC Cys	GAG Glu	GAT Asp 255	ACC Thr	768
GTC Val	TTT Phe	AAG Lys	GGT Gly 260	GTC Val	AGC Ser	ACG Thr	TAC Tyr	ACC Thr 265	ACG Thr	GCC Ala	TCT Ser	GGG Gly	CAG Gln 270	CAG Gln	GTG Val	816
GCC Ala	GGC Gly 275	GTC Val	CTG Leu	GAG Glu	ACG Thr	ACG Thr	GAC Asp 280	AGC Ser	GTC Val	ATG Met	AGA Arg	CGG Arg 285	CTG Leu	ATG Met	AAC Asn	864
CTG Leu 290	CTG Leu	GGG Gly	CAA Gln	GTG Val	GAA Glu	AGT Ser 295	GCC Ala	ATG Met	TCC Ser	GGG Gly	CCC Pro 300	GCG Ala	GCC Ala	TAC Tyr	GCC Ala	912
AGC Ser 305	TAC Tyr	GTT Val	GTC Val	AGG Arg	GGT Gly 310	GCC Ala	AAC Asn	CTC Leu	GTC Val	ACC Thr 315	GCC Ala	GTT Val	AGC Ser	TAC Tyr	GGA Gly 320	960
AGG Arg	GCG Ala	ATG Met	AGA Arg	AAC Asn 325	TTT Phe	GAA Glu	CAG Gln	TTT Phe	ATG Met 330	GCA Ala	CGC Arg	ATA Ile	GTG Val	GAC Asp 335	CAT His	1008
CCC Pro	AAC Asn	GCT Ala 340	CTG Leu	CCG Pro	TCT Ser	GTG Val	GAA Glu	GGT Gly 345	GAC Asp	AAG Lys	GCC Ala	GCT Ala 350	CTG Leu	GCG Ala	GAC Asp	1056
GGA Gly	CAC His 355	GAC Asp	GAG Glu	ATT Ile	CAG Gln	AGA Arg	ACC Thr 360	CGC Arg	ATC Ile	GCC Ala	GCC Ala	TCT Ser 365	CTC Leu	GTG Val	AAG Lys	1104
ATA Ile	GGG Gly 370	GAT Asp	AAG Lys	TTT Phe	GTG Val	GCC Ala 375	ATT Ile	GAA Glu	AGT Ser	TTG Leu	CAG Gln 380	CGC Arg	ATG Met	TAC Tyr	AAC Asn	1152
GAG Glu 385	ACT Thr	CAG Gln	TTT Phe	CCC Pro	TGC Cys 390	CCA Pro	CTG Leu	AAC Asn	CGG Arg 395	CGC Arg	ATC Ile	CAG Gln	TAC Tyr	ACC Thr	TAT Tyr 400	1200
TTC Phe	TTC Phe	CCT Pro	GTT Val	GGC Gly 405	CTT Leu	CAC His	CTT Leu	CCC Pro	GTG Val 410	CCC Pro	CGC Arg	TAC Tyr	TCG Ser	ACA Thr 415	TCC Ser	1248
GTC Val	TCA Ser	GTC Val	AGG Arg 420	GGC Gly	GTA Val	GAA Glu	TCC Ser	CCG Pro 425	GCC Ala	ATC Ile	CAG Gln	TCG Ser 430	ACC Thr	GAG Glu	ACG Thr	1296

[illegible]

TGG	GTG	GTT	AAT	AAA	AAC	AAC	GTG	CCT	CTT	TGC	TTC	GGT	TAC	CAA	AAC	1344
Trp	Val	Val	Asn	Lys	Asn	Asn	Val	Pro	Leu	Cys	Phe	Gly	Tyr	Gln	Asn	
		435					440					445				
GCC	CTC	AAA	AGC	ATA	TGC	CAC	CCT	CGA	ATG	CAC	AAC	CCC	ACC	CAG	TCA	1392
Ala	Leu	Lys	Ser	Ile	Cys	His	Pro	Arg	Met	His	Asn	Pro	Thr	Gln	Ser	
		450				455					460					
GCC	CAG	GCA	CTA	AAC	CAA	GCT	TTT	CCC	GAT	CCC	GAC	GGG	GGA	CAT	GGG	1440
Ala	Gln	Ala	Leu	Asn	Gln	Ala	Phe	Pro	Asp	Pro	Asp	Gly	Gly	His	Gly	
					470					475					480	
TAC	GGT	CTC	AGG	TAT	GAG	CAG	ACG	CCA	AAC	ATG	AAC	CTA	TTC	AGA	ACG	1488
Tyr	Gly	Leu	Arg	Tyr	Glu	Gln	Thr	Pro	Asn	Met	Asn	Leu	Phe	Arg	Thr	
				485					490					495		
TTC	CAC	CAG	TAT	TAC	ATG	GGG	AAA	AAC	GTG	GCA	TTT	GTT	CCC	GAT	GTG	1536
Phe	His	Gln	Tyr	Tyr	Met	Gly	Lys	Asn	Val	Ala	Phe	Val	Pro	Asp	Val	
			500					505					510			
GCC	CAA	AAA	GCG	CTC	GTA	ACC	ACG	GAG	GAT	CTA	CTG	CAC	CCA	ACC	TCT	1584
Ala	Gln	Lys	Ala	Leu	Val	Thr	Thr	Glu	Asp	Leu	Leu	His	Pro	Thr	Ser	
		515					520					525				
CAC	CGT	CTC	CTC	AGA	TTG	GAG	GTC	CAC	CCC	TTC	TTT	GAT	TTT	TTT	GTG	1632
His	Arg	Leu	Leu	Arg	Leu	Glu	Val	His	Pro	Phe	Phe	Asp	Phe	Phe	Val	
		530				535					540					
CAC	CCC	TGT	CCT	GGA	GCG	AGA	GGA	TCG	TAC	CGC	GCC	ACC	CAC	AGA	ACA	1680
His	Pro	Cys	Pro	Gly	Ala	Arg	Gly	Ser	Tyr	Arg	Ala	Thr	His	Arg	Thr	
		545			550					555					560	
ATG	GTT	GGA	AAT	ATA	CCA	CAA	CCG	CTC	GCT	CCA	AGG	GAG	TTT	CAG	GAA	1728
Met	Val	Gly	Asn	Ile	Pro	Gln	Pro	Leu	Ala	Pro	Arg	Glu	Phe	Gln	Glu	
				565					570					575		
AGT	AGA	GGG	GCG	CAG	TTC	GAC	GCT	GTG	ACG	AAT	ATG	ACA	CAC	GTC	ATA	1776
Ser	Arg	Gly	Ala	Gln	Phe	Asp	Ala	Val	Thr	Asn	Met	Thr	His	Val	Ile	
			580					585					590			
GAC	CAG	CTA	ACT	ATT	GAC	GTC	ATA	CAG	GAG	ACG	GCA	TTT	GAC	CCC	GCG	1824
Asp	Gln	Leu	Thr	Ile	Asp	Val	Ile	Gln	Glu	Thr	Ala	Phe	Asp	Pro	Ala	
		595					600					605				
TAT	CCC	CTG	TTC	TGC	TAT	GTA	ATC	GAA	GCA	ATG	ATT	CAC	GGA	CAG	GAA	1872
Tyr	Pro	Leu	Phe	Cys	Tyr	Val	Ile	Glu	Ala	Met	Ile	His	Gly	Gln	Glu	
		610				615					620					
GAA	AAA	TTC	GTG	ATG	AAC	ATG	CCC	CTC	ATT	GCC	CTG	GTC	ATT	CAA	ACC	1920
Glu	Lys	Phe	Val	Met	Asn	Met	Pro	Leu	Ile	Ala	Leu	Val	Ile	Gln	Thr	
					630					635					640	
TAC	TGG	GTC	AAC	TCG												

ATC ACA CAT CAT GTT TCG GTC TCG GTC CTC CTC GAC CCG CAT CTG CTG CCT CCC Ile Thr His Leu Val Ser Ala Leu Leu Asp Pro His Leu Leu Pro Pro 705 710 715 720	2160
TTT GCC TAC CAC GAT GTC TTT ACG GAT CTT ATG CAG AAG TCA TCC AGA Phe Ala Tyr His Asp Val Phe Thr Asp Leu Met Gln Lys Ser Ser Arg 725 730 735	2208
CAA CCC ATA ATC AAG ATC GGG GAT CAA AAC TAC GAC AAC CCT CAA AAT Gln Pro Ile Ile Lys Ile Gly Asp Gln Asn Tyr Asp Asn Pro Gln Asn 740 745 750	2256
AGG GCG ACA TTC ATC AAC CTC AGG GGT CGC ATG GAG GAC CTA GTC AAT Arg Ala Thr Phe Ile Asn Leu Arg Gly Arg Met Glu Asp Leu Val Asn 755 760 765	2304
AAC CTT GTT AAC ATT TAC CAG ACA AGG GTC AAT GAG GAC CAT GAC GAG Asn Leu Val Asn Ile Tyr Gln Thr Arg Val Asn Glu Asp His Asp Glu 770 775 780	2352
AGA CAC GTC CTG GAC GTG GCG CCC CTG GAC GAG AAT GAC TAC AAC CCG Arg His Val Leu Asp Val Ala Pro Leu Asp Glu Asn Asp Tyr Asn Pro 785 790 795 800	2400
GTC CTC GAG AAG CTA TTC TAC TAT GTT TTA ATG CCG GTG TGC AGT AAC Val Leu Glu Lys Leu Phe Tyr Tyr Val Leu Met Pro Val Cys Ser Asn 805 810 815	2448
GGC CAC ATG TGC GGT ATG GGG GTC GAC TAT CAA AAC GTG GCC CTG ACG Gly His Met Cys Gly Met Gly Val Asp Tyr Gln Asn Val Ala Leu Thr 820 825 830	2496
CTG ACT TAC AAC GGC CCC GTC TTT GCG GAC GTC GTG AAC GCA CAG GAT Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Ala Gln Asp 835 840 845	2544
GAT ATT CTA CTG CAC CTG GAG AAC GGA ACC TTG AAG GAC ATT CTG CAG Asp Ile Leu Leu His Leu Glu Asn Gly Thr Leu Lys Asp Ile Leu Gln 850 855 860	2592
GCA GGC GAC ATA CGC CCG ACG GTG GAC ATG ATC AGG GTG CTG TGC ACC Ala Gly Asp Ile Arg Pro Thr Val Asp Met Ile Arg Val Leu Cys Thr 865 870 875 880	2640
TCG TTT CTG ACG TGC CCT TTC GTC ACC CAG GCC GCT CGC GTG ATC ACA Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ala Arg Val Ile Thr 885 890 895	2688
AAG CGG GAC CCG GCC CAG AGT TTT GCC ACG CAC GAA TAC GGG AAG GAT Lys Arg Asp Pro Ala Gln Ser Phe Ala Thr His Glu Tyr Gly Lys Asp 900 905 910	2736
GTG GCG CAG ACC GTG CTT GTT AAT GGC TTT GGT GCG TTC GCG GTG GCG Val Ala Gln Thr Val Leu Val Asn Gly Phe Gly Ala Phe Ala Val Ala 915 920 925	2784
GAC CGC TCT CGC GAG GCG GCG GAG ACT ATG TTT TAT CCG GTA CCC TTT Asp Arg Ser Arg Glu Ala Ala Glu Thr Met Phe Tyr Pro Val Pro Phe 930 935 940	2832
AAC AAG CTC TAC GCT GAC CCG TTG GTG GCT GCC ACA CTG CAT CCG CTC Asn Lys Leu Tyr Ala Asp Pro Leu Val Ala Thr Leu His Pro Leu 945 950 955 960	2880
CTG CCA AAC TAT GTC ACC AGG CTC CCC AAC CAG AGA AAC GCG GTG GTC Leu Pro Asn Tyr Val Thr Arg Leu Pro Asn Gln Arg Asn Ala Val Val 965 970 975	2928

TTT	AAc	GTG	CCA	TCC	AAT	CTC	ATG	GCA	GAA	TAT	GAG	GAA	TGG	CAC	AAG	2976
Phe	Asn	Val	Pro	Ser	Asn	Leu	Met	Ala	Glu	Tyr	Glu	Glu	Trp	His	Lys	
		980						985					990			
TCG	CCC	GTC	GCG	GCG	TAT	GCC	GCG	TCT	TGT	CAG	GCC	ACC	CCG	GGC	GCC	3024
Ser	Pro	Val	Ala	Ala	Tyr	Ala	Ala	Ser	Cys	Gln	Ala	Thr	Pro	Gly	Ala	
		995					1000					1005				
ATT	AGC	GCC	ATG	GTG	AGC	ATG	CAC	CAA	AAA	CTA	TCT	GCC	CCC	AGT	TTC	3072
Ile	Ser	Ala	Met	Val	Ser	Met	His	Gln	Lys	Leu	Ser	Ala	Pro	Ser	Phe	
	1010					1015					1020					
ATT	TGC	CAG	GCA	AAA	CAC	CGC	ATG	CAC	CCT	GGT	TTT	GCC	ATG	ACA	GTC	3120
Ile	Cys	Gln	Ala	Lys	His	Arg	Met	His	Pro	Gly	Phe	Ala	Met	Thr	Val	
	1025				1030					1035					1040	
GTC	AGG	ACG	GAC	GAG	GTT	CTA	GCA	GAG	CAC	ATC	CTA	TAC	TGC	TCC	AGG	3168
Val	Arg	Thr	Asp	Glu	Val	Leu	Ala	Glu	His	Ile	Leu	Tyr	Cys	Ser	Arg	
				1045					1050					1055		
GCG	TCG	ACA	TCC	ATG	TTT	GTG	GGC	TTG	CCT	TCG	GTG	GTA	CGG	CGC	GAG	3216
Ala	Ser	Thr	Ser	Met	Phe	Val	Gly	Leu	Pro	Ser	Val	Val	Arg	Arg	Glu	
			1060					1065					1070			
GTA	CGT	TCG	GAC	GCG	GTG	ACT	TTT	GAA	ATT	ACC	CAC	GAG	ATC	GCT	TCC	3264
Val	Arg	Ser	Asp	Ala	Val	Thr	Phe	Glu	Ile	Thr	His	Glu	Ile	Ala	Ser	
		1075					1080					1085				
CTG	CAC	ACC	GCA	CTT	GGC	TAC	TCA	TCA	GTC	ATC	GCC	CCG	GCC	CAC	GTG	3312
Leu	His	Thr	Ala	Leu	Gly	Tyr	Ser	Ser	Val	Ile	Ala	Pro	Ala	His	Val	
	1090					1095					1100					
GCC	GCC	ATA	ACT	ACA	GAC	ATG	GGA	GTA	CAT	TGT	CAG	GAC	CTC	TTT	ATG	3360
Ala	Ala	Ile	Thr	Thr	Asp	Met	Gly	Val	His	Cys	Gln	Asp	Leu	Phe	Met	
	1105				1110					1115					1120	
ATT	TTC	CCA	GGG	GAC	GCG	TAT	CAG	GAC	CGC	CAG	CTG	CAT	GAC	TAT	ATC	3408
Ile	Phe	Pro	Gly	Asp	Ala	Tyr	Gln	Asp	Arg	Gln	Leu	His	Asp	Tyr	Ile	
				1125					1130					1135		
AAA	ATG	AAA	GCG	GGC	GTG	CAA	ACC	GGC	TCA	CCG	GGA	AAC	AGA	ATG	GAT	3456
Lys	Met	Lys	Ala	Gly	Val	Gln	Thr	Gly	Ser	Pro	Gly	Asn	Arg	Met	Asp	
		1140						1145					1150			
CAC	GTG	GGA	TAC	ACT	GCT	GGG	GTT	CCT	CGC	TGC	GAG	AAC	CTG	CCC	GGT	3504
His	Val	Gly	Tyr	Thr	Ala	Gly	Val	Pro	Arg	Cys	Glu	Asn	Leu	Pro	Gly	
		1155					1160					1165				
TTG	AGT	CAT	GGT	CAG	CTG	GCA	ACC	TGC	GAG	ATA	ATT	CCC	ACG	CCG	GTC	3552
Leu	Ser	His	Gly	Gln	Leu	Ala	Thr	Cys	Glu	Ile	Ile	Pro	Thr	Pro	Val	
	1170					1175					1180					

193

CTA TAC AAT ATC ACC TTT CGC CAG ACT GCG CTG CCG GGC ATG TAC AGT 3792
 Leu Tyr Asn Ile Thr Phe Arg Gln Thr Ala Leu Pro Gly Met Tyr Ser
 1250 1255 1260

CCT TGT CGG CAG TTC TTC CAC AAG GAA GAC ATT ATG CGG TAC AAT AGG 3840
 Pro Cys Arg Gln Phe Phe His Lys Glu Asp Ile Met Arg Tyr Asn Arg
 1265 1270 1275 1280

GGG TTG TAC ACT TTG GTT AAT GAG TAT TCT GCC AGG CTT GCT GGG GCC 3888
 Gly Leu Tyr Thr Leu Val Asn Glu Tyr Ser Ala Arg Leu Ala Gly Ala
 1285 1290 1295

CCC GCC ACC AGC ACT ACA GAC CTC CAG TAC GTC GTG GTC AAC GGT ACA 3936
 Pro Ala Thr Ser Thr Thr Asp Leu Gln Tyr Val Val Val Asn Gly Thr
 1300 1305 1310

GAC GTG TTT TTG GAC CAG CCT TGC CAT ATG CTG CAG GAG GCC TAT CCC 3984
 Asp Val Phe Leu Asp Gln Pro Cys His Met Leu Gln Glu Ala Tyr Pro
 1315 1320 1325

ACG CTC GCC GCC AGC CAC AGA GTT ATG CTT GCC GAG TAC ATG TCA AAC 4032
 Thr Leu Ala Ala Ser His Arg Val Met Leu Ala Glu Tyr Met Ser Asn
 1330 1335 1340

AAG CAG ACA CAC GCC CCA GTA CAC ATG GGC CAG TAT CTC ATT GAA GAG 4080
 Lys Gln Thr His Ala Pro Val His Met Gly Gln Tyr Leu Ile Glu Glu
 1345 1350 1355 1360

GTG GCG CCG ATG AAG AGA CTA TTA AAG CTC GGA AAC AAG GTG GTG TAT 4128
 Val Ala Pro Met Lys Arg Leu Leu Lys Leu Gly Asn Lys Val Val Tyr
 1365 1370 1375

TAG 4131

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1376 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Glu Ala Thr Leu Glu Gln Arg Pro Phe Pro Tyr Leu Ala Thr Glu
 1 5 10 15

Ala Asn Leu Leu Thr Gln Ile Lys Glu Ser Ala Ala Asp Gly Leu Phe
 20 25 30

Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Ser Val
 35 40 45

Arg Phe Glu Ala Leu Leu Gly Val Tyr Thr Asn Val Val Glu Phe Val
 50 55 60

Lys Phe Leu Glu Thr Ala Leu Ala Ala Ala Cys Val Asn Thr Glu Phe
 65 70 75 80

Lys Asp Leu Arg Arg Met Ile Asp Gly Lys Ile Gln Phe Lys Ile Ser
 85 90 95

Met Pro Thr Ile Ala His Gly Asp Gly Arg Arg Pro Asn Lys Glu Arg
 100 105 110

Gln Tyr Ile Val Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125
Ile Glu Leu Ala Ala Ala Asp Ile Glu Leu Leu Phe Ala Glu Lys Glu
130 135 140
Thr Pro Leu Asp Phe Thr Glu Tyr Ala Gly Ala Ile Lys Thr Ile Thr
145 150 155 160
Ser Ala Leu Gln Phe Gly Met Asp Ala Leu Glu Arg Gly Leu Val Asp
165 170 175
Thr Val Leu Ala Val Lys Leu Arg His Ala Pro Pro Val Phe Ile Leu
180 185 190
Lys Thr Leu Gly Asp Pro Val Tyr Ser Glu Arg Gly Leu Lys Lys Ala
195 200 205
Val Lys Ser Asp Met Val Ser Met Phe Lys Ala His Leu Ile Glu His
210 215 220
Ser Phe Phe Leu Asp Lys Ala Glu Leu Met Thr Arg Gly Lys Gln Tyr
225 230 235 240
Val Leu Thr Met Leu Ser Asp Met Leu Ala Ala Val Cys Glu Asp Thr
245 250 255
Val Phe Lys Gly Val Ser Thr Tyr Thr Thr Ala Ser Gly Gln Gln Val
260 265 270
Ala Gly Val Leu Glu Thr Thr Asp Ser Val Met Arg Arg Leu Met Asn
275 280 285
Leu Leu Gly Gln Val Glu Ser Ala Met Ser Gly Pro Ala Ala Tyr Ala
290 295 300
Ser Tyr Val Val Arg Gly Ala Asn Leu Val Thr Ala Val Ser Tyr Gly
305 310 315 320
Arg Ala Met Arg Asn Phe Glu Gln Phe Met Ala Arg Ile Val Asp His
325 330 335
Pro Asn Ala Leu Pro Ser Val Glu Gly Asp Lys Ala Ala Leu Ala Asp
340 345 350
Gly His Asp Glu Ile Gln Arg Thr Arg Ile Ala Ala Ser Leu Val Lys
355 360 365
Ile Gly Asp Lys Phe Val Ala Ile Glu Ser Leu Gln Arg Met Tyr Asn
370 375 380
Glu Thr Gln Phe Pro Cys Pro Leu Asn Arg Arg Ile Gln Tyr Thr Tyr
385 390 395 400
Phe Phe Pro Val Gly Leu His Leu Pro Val Pro Arg Tyr Ser Thr Ser
405 410 415
Val Ser Val Arg Gly Val Glu Ser Pro Ala Ile Gln Ser Thr Glu Thr
420 425 430
Trp Val Val Asn Lys Asn Asn Val Pro Leu Cys Phe Gly Tyr Gln Asn
435 440 445
Ala Leu Lys Ser Ile Cys His Pro Arg Met His Asn Pro Thr Gln Ser
450 455 460
Ala Gln Ala Leu Asn Gln Ala Phe Pro Asp Pro Asp Gly Gly His Gly

WO 96/06159

195

465 470 475 480
 Tyr Gly Leu Arg Tyr Glu Gln Thr Pro Asn Met Asn Leu Phe Arg Thr
 485 490 495
 Phe His Gln Tyr Tyr Met Gly Lys Asn Val Ala Phe Val Pro Asp Val
 500 505 510
 Ala Gln Lys Ala Leu Val Thr Thr Glu Asp Leu Leu His Pro Thr Ser
 515 520 525
 His Arg Leu Leu Arg Leu Glu Val His Pro Phe Phe Asp Phe Phe Val
 530 535 540
 His Pro Cys Pro Gly Ala Arg Gly Ser Tyr Arg Ala Thr His Arg Thr
 545 550 555 560
 Met Val Gly Asn Ile Pro Gln Pro Leu Ala Pro Arg Glu Phe Gln Glu
 565 570 575
 Ser Arg Gly Ala Gln Phe Asp Ala Val Thr Asn Met Thr His Val Ile
 580 585 590
 Asp Gln Leu Thr Ile Asp Val Ile Gln Glu Thr Ala Phe Asp Pro Ala
 595 600 605
 Tyr Pro Leu Phe Cys Tyr Val Ile Glu Ala Met Ile His Gly Gln Glu
 610 615 620
 Glu Lys Phe Val Met Asn Met Pro Leu Ile Ala Leu Val Ile Gln Thr
 625 630 635 640
 Tyr Trp Val Asn Ser Gly Lys Leu Ala Phe Val Asn Ser Tyr His Met
 645 650 655
 Val Arg Phe Ile Cys Thr His Ile Gly Asn Gly Ser Ile Pro Lys Glu
 660 665 670
 Ala His Gly His Tyr Arg Lys Ile Leu Gly Glu Leu Ile Ala Leu Glu
 675 680 685
 Gln Ala Leu Leu Lys Leu Ala Gly His Glu Thr Val Gly Arg Thr Pro
 690 695 700
 Ile Thr His Leu Val Ser Ala Leu Leu Asp Pro His Leu Leu Pro Pro
 705 710 715 720
 Phe Ala Tyr His Asp Val Phe Thr Asp Leu Met Gln Lys Ser Ser Arg
 725 730 735
 Gln Pro Ile Ile Lys Ile Gly Asp Gln Asn Tyr Asp Asn Pro Gln Asn
 740 745 750
 Arg Ala Thr Phe Ile Asn Leu Arg Gly Arg Met Glu Asp Leu Val Asn
 755 760 765
 Asn Leu Val Asn Ile Tyr Gln Thr Arg Val Asn Glu Asp His Asp Glu
 770 775 780
 Arg His Val Leu Asp Val Ala Pro Leu Asp Glu Asn Asp Tyr Asn Pro
 785 790 795 800
 Val Leu Glu Lys Leu Phe Tyr Tyr Val Leu Met Pro Val Cys Ser Asn
 805 810 815
 Gly His Met Cys Gly Met Gly Val Asp Tyr Gln Asn Val Ala Leu Thr
 820 825 830

196

Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Ala Gln Asp
 835 840 845
 Asp Ile Leu Leu His Leu Glu Asn Gly Thr Leu Lys Asp Ile Leu Gln
 850 855 860
 Ala Gly Asp Ile Arg Pro Thr Val Asp Met Ile Arg Val Leu Cys Thr
 865 870 875 880
 Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ala Arg Val Ile Thr
 885 890 895
 Lys Arg Asp Pro Ala Gln Ser Phe Ala Thr His Glu Tyr Gly Lys Asp
 900 905 910
 Val Ala Gln Thr Val Leu Val Asn Gly Phe Gly Ala Phe Ala Val Ala
 915 920 925
 Asp Arg Ser Arg Glu Ala Ala Glu Thr Met Phe Tyr Pro Val Pro Phe
 930 935 940
 Asn Lys Leu Tyr Ala Asp Pro Leu Val Ala Ala Thr Leu His Pro Leu
 945 950 955 960
 Leu Pro Asn Tyr Val Thr Arg Leu Pro Asn Gln Arg Asn Ala Val Val
 965 970 975
 Phe Asn Val Pro Ser Asn Leu Met Ala Glu Tyr Glu Glu Trp His Lys
 980 985 990
 Ser Pro Val Ala Ala Tyr Ala Ala Ser Cys Gln Ala Thr Pro Gly Ala
 995 1000 1005
 Ile Ser Ala Met Val Ser Met His Gln Lys Leu Ser Ala Pro Ser Phe
 1010 1015 1020
 Ile Cys Gln Ala Lys His Arg Met His Pro Gly Phe Ala Met Thr Val
 1025 1030 1035 1040
 Val Arg Thr Asp Glu Val Leu Ala Glu His Ile Leu Tyr Cys Ser Arg
 1045 1050 1055
 Ala Ser Thr Ser Met Phe Val Gly Leu Pro Ser Val Val Arg Arg Glu
 1060 1065 1070
 Val Arg Ser Asp Ala Val Thr Phe Glu Ile Thr His Glu Ile Ala Ser
 1075 1080 1085
 Leu His Thr Ala Leu Gly Tyr Ser Ser Val Ile Ala Pro Ala His Val
 1090 1095 1100
 Ala Ala Ile Thr Thr Asp Met Gly Val His Cys Gln Asp Leu Phe Met
 1105 1110 1115 1120
 Ile Phe Pro Gly Asp Ala Tyr Gln Asp Arg Gln Leu His Asp Tyr Ile
 1125 1130 1135
 Lys Met Lys Ala Gly Val Gln Thr Gly Ser Pro Gly Asn Arg Met Asp
 1140 1145 1150
 His Val Gly Tyr Thr Ala Gly Val Pro Arg Cys Glu Asn Leu Pro Gly
 1155 1160 1165
 Leu Ser His Gly Gln Leu Ala Thr Cys Glu Ile Ile Pro Thr Pro Val
 1170 1175 1180
 Thr Ser Asp Val Ala Tyr Phe Gln Thr Pro Ser Asn Pro Arg Gly Arg

006290-6242960

197

1185 1190 1195 1200
 Ala Ala Ser Val Val Ser Cys Asp Ala Tyr Ser Asn Glu Ser Ala Glu
 1205 1210 1215
 Arg Leu Phe Tyr Asp His Ser Ile Pro Asp Pro Ala Tyr Glu Cys Arg
 1220 1225 1230
 Ser Thr Asn Asn Pro Trp Ala Ser Gln Arg Gly Ser Leu Gly Asp Val
 1235 1240 1245
 Leu Tyr Asn Ile Thr Phe Arg Gln Thr Ala Leu Pro Gly Met Tyr Ser
 1250 1255 1260
 Pro Cys Arg Gln Phe Phe His Lys Glu Asp Ile Met Arg Tyr Asn Arg
 1265 1270 1275 1280
 Gly Leu Tyr Thr Leu Val Asn Glu Tyr Ser Ala Arg Leu Ala Gly Ala
 1285 1290 1295
 Pro Ala Thr Ser Thr Thr Asp Leu Gln Tyr Val Val Val Asn Gly Thr
 1300 1305 1310
 Asp Val Phe Leu Asp Gln Pro Cys His Met Leu Gln Glu Ala Tyr Pro
 1315 1320 1325
 Thr Leu Ala Ala Ser His Arg Val Met Leu Ala Glu Tyr Met Ser Asn
 1330 1335 1340
 Lys Gln Thr His Ala Pro Val His Met Gly Gln Tyr Leu Ile Glu Glu
 1345 1350 1355 1360
 Val Ala Pro Met Lys Arg Leu Leu Lys Leu Gly Asn Lys Val Val Tyr
 1365 1370 1375

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1143
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGC ATT CGG GGA CAG ACC TTT AAC CTG CTC TAC GTA GAC GAG GCG AAT 48
 Ser Ile Arg Gly Gln Thr Phe Asn Leu Leu Tyr Val Asp Glu Ala Asn
 1 5 10 15
 TTT ATT AAA AAG GAT GCA CTG CCG GCT ATT CTG GGT TTC ATG CTT CAG 96
 Phe Ile Lys Lys Asp Ala Leu Pro Ala Ile Leu Gly Phe Met Leu Gln
 20 25 30

198

AAA	GAC	GCC	AAG	CTT	ATA	TTT	ATA	TCA	TCC	GTG	AAC	TCG	TCA	GAC	CGC	144
Lys	Asp	Ala	Lys	Leu	Ile	Phe	Ile	Ser	Ser	Val	Asn	Ser	Ser	Asp	Arg	
		35					40					45				
TCC	ACG	AGT	TTC	CTG	CTT	AAC	CTC	AGG	AAC	GCC	CAG	GAA	AAG	ATG	CTG	192
Ser	Thr	Ser	Phe	Leu	Leu	Asn	Leu	Arg	Asn	Ala	Gln	Glu	Lys	Met	Leu	
	50					55					60					
AAT	GTG	GTC	AGT	TAC	GTG	TGT	GCG	GAC	CAC	CGA	GAA	GAT	TTC	CAC	CTG	240
Asn	Val	Val	Ser	Tyr	Val	Cys	Ala	Asp	His	Arg	Glu	Asp	Phe	His	Leu	
65					70					75					80	
CAA	GAC	GCA	CTA	GTG	TCC	TGT	CCT	TGT	TAC	AGA	CTG	CAC	ATT	CCG	ACG	288
Gln	Asp	Ala	Leu	Val	Ser	Cys	Pro	Cys	Tyr	Arg	Leu	His	Ile	Pro	Thr	
				85					90					95		
TAC	ATC	ACC	ATC	GAC	GAA	TCC	ATC	AAA	ACC	ACC	ACC	AAC	CTC	TTT	ATG	336
Tyr	Ile	Thr	Ile	Asp	Glu	Ser	Ile	Lys	Thr	Thr	Thr	Asn	Leu	Phe	Met	
			100					105					110			
GAG	GGG	GCA	TTC	GAC	ACC	GAA	CTA	ATG	GGC	GAG	GGA	GCA	GCG	TCG	TCA	384
Glu	Gly	Ala	Phe	Asp	Thr	Glu	Leu	Met	Gly	Glu	Gly	Ala	Ala	Ser	Ser	
	115						120					125				
AAT	GCT	ACG	CTT	TAC	CGC	GTG	GTG	GGT	GAC	GCA	GCG	CTG	ACA	CAG	TTT	432
Asn	Ala	Thr	Leu	Tyr	Arg	Val	Val	Gly	Asp	Ala	Ala	Leu	Thr	Gln	Phe	
130						135					140					
GAC	ATG	TGT	CGG	GTA	GAC	ACC	ACC	GCC	CAG	GAG	GTT	CAG	AAG	TGC	CTT	480
Asp	Met	Cys	Arg	Val	Asp	Thr	Thr	Ala	Gln	Glu	Val	Gln	Lys	Cys	Leu	
145					150					155					160	
GGA	AAA	CAG	CTG	TTT	GTT	TAC	ATC	GAC	CCC	GCG	TAT	ACG	AAC	AAC	ACG	528
Gly	Lys	Gln	Leu	Phe	Val	Tyr	Ile	Asp	Pro	Ala	Tyr	Thr	Asn	Asn	Thr	
			165						170					175		
GAG	GCG	TCC	GGT	ACT	GGC	GTG	GGC	GCC	GTT	GTC	ACG	AGT	ACT	CAG	ACT	576
Glu	Ala	Ser	Gly	Thr	Gly	Val	Gly	Ala	Val	Val	Thr	Ser	Thr	Gln	Thr	
			180					185						190		
CCC	ACC	AGA	AGC	CTC	ATA	TTG	GGC	ATG	GAG	CAT	TTC	TTC	CTG	CGC	GAC	624
Pro	Thr	Arg	Ser	Leu	Ile	Leu	Gly	Met	Glu	His	Phe	Phe	Leu	Arg	Asp	
		195					200					205				
CTC	ACT	GGC	GCA	GCT	GCT	TAC	GAG	ATA	GCG	TCC	TGC	GCA	TGC	ACG	ATG	672
Leu	Thr	Gly	Ala	Ala	Ala	Tyr	Glu	Ile	Ala	Ser	Cys	Ala	Cys	Thr	Met	
	210					215					220					
ATT	AAG	GCG	ATC	GCT	GTG	CTC	CAC	ACC	ACA	ATT	GAG	CGC	GTG	AAC	GCG	720
Ile	Lys	Ala	Ile	Ala	Val	Leu	His	Thr	Thr	Ile	Glu	Arg	Val	Asn	Ala	
225					230					235					240	
GCG	GTC	GAA	GGC	AAC	AGC	AGC	CAA	GAT	TCT	GGG	GTG	GCC	ATT	GCA	ACC	768
Ala	Val	Glu	Gly	Asn	Ser	Ser	Gln	Asp	Ser	Gly	Val	Ala	Ile	Ala	Thr	
				245				250						255		
GTC	CTT	AAC	GAA	ATA	TGC	CCG	CTC	CCC	ATA	CAT	TTT	CTA	CAC	TAT	ACT	816
Val	Leu	Asn	Glu	Ile	Cys	Pro	Leu	Pro	Ile	His	Phe	Leu	His	Tyr	Thr	
			260					265					270			
GAC	AAG	AGC	AGC	GCC	CTG	CAG	TGG	CCA	ATT	TAC	ATG	TTG	GGA	GGC	GAG	864
Asp	Lys	Ser	Ser	Ala	Leu	Gln	Trp	Pro	Ile	Tyr	Met	Leu	Gly	Gly	Glu	
		275					280					285				
AAA	TCC	TCC	GCG	TTT	GAG	ACA	TTC	ATC	TAC	GCT	CTG	AAC	TCC	GGC	ACC	912
Lys	Ser	Ser	Ala	Phe	Glu	Thr	Phe	Ile	Tyr	Ala	Leu	Asn	Ser	Gly	Thr	
						295					300					
	290															

006290.06290

199

CTG AGC GCC AGC CAG ACG GTG GTG TCC AAC ACC ATC AAA ATA TCA TTT 960
 Leu Ser Ala Ser Gln Thr Val Val Ser Asn Thr Ile Lys Ile Ser Phe
 305 310 315 320

GAC CCG GTG ACC TAC CTG GTA GAA CAG GTC CGC GCG ATC AAG TGC GTC 1008
 Asp Pro Val Thr Tyr Leu Val Glu Gln Val Arg Ala Ile Lys Cys Val
 325 330 335

CCG CTT AGG GAT GGA GGG CAG TCA TAC AGC GCC AAG CAA AAG CAC ATG 1056
 Pro Leu Arg Asp Gly Gly Gln Ser Tyr Ser Ala Lys Gln Lys His Met
 340 345 350

TCG GAC GAC TTA CTT GTG GCA GTT GTC ATG GCC CAT TTT ATG GCT ACC 1104
 Ser Asp Asp Leu Leu Val Ala Val Val Met Ala His Phe Met Ala Thr
 355 360 365

GAT GAT AGA CAC ATG TAC AAG CCC ATA TCC CCA CAA TAA 1143
 Asp Asp Arg His Met Tyr Lys Pro Ile Ser Pro Gln
 370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ser Ile Arg Gly Gln Thr Phe Asn Leu Leu Tyr Val Asp Glu Ala Asn
 1 5 10 15

Phe Ile Lys Lys Asp Ala Leu Pro Ala Ile Leu Gly Phe Met Leu Gln
 20 25 30

Lys Asp Ala Lys Leu Ile Phe Ile Ser Ser Val Asn Ser Ser Asp Arg
 35 40 45

Ser Thr Ser Phe Leu Leu Asn Leu Arg Asn Ala Gln Glu Lys Met Leu
 50 55 60

Asn Val Val Ser Tyr Val Cys Ala Asp His Arg Glu Asp Phe His Leu
 65 70 75 80

Gln Asp Ala Leu Val Ser Cys Pro Cys Tyr Arg Leu His Ile Pro Thr
 85 90 95

Tyr Ile Thr Ile Asp Glu Ser Ile Lys Thr Thr Thr Asn Leu Phe Met
 100 105 110

Glu Gly Ala Phe Asp Thr Glu Leu Met Gly Glu Gly Ala Ala Ser Ser
 115 120 125

Asn Ala Thr Leu Tyr Arg Val Val Gly Asp Ala Ala Leu Thr Gln Phe
 130 135 140

Asp Met Cys Arg Val Asp Thr Thr Ala Gln Glu Val Gln Lys Cys Leu
 145 150 155 160

Gly Lys Gln Leu Phe Val Tyr Ile Asp Pro Ala Tyr Thr Asn Asn Thr
 165 170 175

Glu Ala Ser Gly Thr Gly Val Gly Ala Val Val Thr Ser Thr Gln Thr
 180 185 190

0060749.062900

200

Pro Thr Arg Ser Leu Ile Leu Gly Met Glu His Phe Phe Leu Arg Asp
 195 200 205

Leu Thr Gly Ala Ala Ala Tyr Glu Ile Ala Ser Cys Ala Cys Thr Met
 210 215 220

Ile Lys Ala Ile Ala Val Leu His Thr Thr Ile Glu Arg Val Asn Ala
 225 230 235 240

Ala Val Glu Gly Asn Ser Ser Gln Asp Ser Gly Val Ala Ile Ala Thr
 245 250 255

Val Leu Asn Glu Ile Cys Pro Leu Pro Ile His Phe Leu His Tyr Thr
 260 265 270

Asp Lys Ser Ser Ala Leu Gln Trp Pro Ile Tyr Met Leu Gly Gly Glu
 275 280 285

Lys Ser Ser Ala Phe Glu Thr Phe Ile Tyr Ala Leu Asn Ser Gly Thr
 290 295 300

Leu Ser Ala Ser Gln Thr Val Val Ser Asn Thr Ile Lys Ile Ser Phe
 305 310 315 320

Asp Pro Val Thr Tyr Leu Val Glu Gln Val Arg Ala Ile Lys Cys Val
 325 330 335

Pro Leu Arg Asp Gly Gly Gln Ser Tyr Ser Ala Lys Gln Lys His Met
 340 345 350

Ser Asp Asp Leu Leu Val Ala Val Val Met Ala His Phe Met Ala Thr
 355 360 365

Asp Asp Arg His Met Tyr Lys Pro Ile Ser Pro Gln
 370 375 380

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..234
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG GGT GAG CCA GTG GAT CCT GGA CAT GTG GTG AAT GAG AAA GAT TTT
 Met Gly Glu Pro Val Asp Pro Gly His Val Val Asn Glu Lys Asp Phe
 1 5 10 15

GAG GAG TGT GAA CAA TTT TTC AGT CAA CCC CTT AGG GAG CAA GTG GTC
 Glu Glu Cys Glu Gln Phe Phe Ser Gln Pro Leu Arg Glu Gln Val Val
 20 25 30

48

96

006290" 6442990

[illegible]

(2) INFORMATION FOR SEQ ID NO:7:

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Gly	Glu	Pro	Val	Asp	Pro	Gly	His	Val	Val	Asn	Glu	Lys	Asp	Phe
1				5					10					15	
Glu	Glu	Cys	Glu	Gln	Phe	Phe	Ser	Gln	Pro	Leu	Arg	Glu	Gln	Val	Val
			20					25					30		
Ala	Gly	Val	Arg	Ala	Leu	Asp	Gly	Leu	Gly	Leu	Ala	Asp	Ser	Leu	Cys
		35					40					45			
His	Lys	Thr	Glu	Arg	Leu	Cys	Leu	Leu	Met	Asp	Leu	Val	Gly	Thr	Glu
	50					55					60				
Cys	Phe	Ala	Arg	Val	Cys	Arg	Leu	Asp	Thr	Gly	Ala	Lys			
65					70					75					

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 585 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) **FEATURE:**

(A) NAME/KEY: CDS
(B) LOCATION: 1..585
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATG	AAG	AGT	GTG	GCG	AGT	CCC	TTA	TGT	CAG	TTC	CAC	GGC	GTG	TTT	TGC	48
Met	Lys	Ser	Val	Ala	Ser	Pro	Leu	Cys	Gln	Phe	His	Gly	Val	Phe	Cys	
1				5				10						15		
CTG	TAC	CAG	TGT	CGC	CAG	TGC	CTG	GCA	TAC	CAC	GTG	TGT	GAT	GGG	GGC	96

202

Leu	Tyr	Gln	Cys	Arg	Gln	Cys	Leu	Ala	Tyr	His	Val	Cys	Asp	Gly	Gly		
			20					25					30				
GCC	GAA	TGC	GTT	CTC	CTG	CAT	ACG	CCG	GAG	AGC	GTC	ATC	TGC	GAA	CTA		144
Ala	Glu	Cys	Val	Leu	Leu	His	Thr	Pro	Glu	Ser	Val	Ile	Cys	Glu	Leu		
		35					40					45					
ACG	GGT	AAC	TGC	ATG	CTC	GGC	AAC	ATT	CAA	GAG	GGC	CAG	TTT	TTA	GGG		192
Thr	Gly	Asn	Cys	Met	Leu		Asn	Ile	Gln	Glu	Gly	Gln	Phe	Leu	Gly		
	50					55					60						
CCG	GTA	CCG	TAT	CGG	ACT	TTG	GAT	AAC	CAG	GTT	GAC	AGG	GAC	GCA	TAT		240
Pro	Val	Pro	Tyr	Arg	Thr	Leu	Asp	Asn	Gln	Val	Asp	Arg	Asp	Ala	Tyr		
	65				70					75					80		
CAC	GGG	ATG	CTA	GCG	TGT	CTG	AAA	CGG	GAC	ATT	GTG	CGG	TAT	TTG	CAG		288
His	Gly	Met	Leu	Ala	Cys	Leu	Lys	Arg	Asp	Ile	Val	Arg	Tyr	Leu	Gln		
				85					90					95			
ACA	TGG	CCG	GAC	ACC	ACC	GTA	ATC	GTG	CAG	GAA	ATA	GCC	CTG	GGG	GAC		336
Thr	Trp	Pro	Asp	Thr	Thr	Val	Ile	Val	Gln	Glu	Ile	Ala	Leu	Gly	Asp		
			100					105					110				
GGC	GTC	ACC	GAC	ACC	ATC	TCG	GCC	ATT	ATA	GAT	GAA	ACA	TTC	GGT	GAG		384
Gly	Val	Thr	Asp	Thr	Ile	Ser	Ala	Ile	Ile	Asp	Glu	Thr	Phe	Gly	Glu		
		115					120					125					
TGT	CTT	CCC	GTA	CTG	GGG	GAG	GCC	CAA	GGC	GGG	TAC	GCC	CTG	GTC	TGT		432
Cys	Leu	Pro	Val	Leu	Gly	Glu	Ala	Gln	Gly	Gly	Tyr	Ala	Leu	Val	Cys		
	130					135					140						
AGC	ATG	TAT	CTG	CAC	GTT	ATC	GTC	TCC	ATC	TAT	TCG	ACA	AAA	ACG	GTG		480
Ser	Met	Tyr	Leu	His	Val	Ile	Val	Ser	Ile	Tyr	Ser	Thr	Lys	Thr	Val		
	145				150					155					160		
TAC	AAC	AGT	ATG	CTA	TTT	AAA	TGC	ACA	AAG	AAT	AAA	AAG	TAC	GAC	TGC		528
Tyr	Asn	Ser	Met	Leu	Phe	Lys	Cys	Thr	Lys	Asn	Lys	Lys	Tyr	Asp	Cys		
				165					170					175			
ATT	GCC	AAG	CGG	GTG	CGG	ACA	AAA	TGG	ATG	CGC	ATG	CTA	TCA	ACG	AAA		576
Ile	Ala	Lys	Arg	Val	Arg	Thr	Lys	Trp	Met	Arg	Met	Leu	Ser	Thr	Lys		
			180					185					190				
GAT	ACG	TAG															585
Asp	Thr																
			195														

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 194 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Lys	Ser	Val	Ala	Ser	Pro	Leu	Cys	Gln	Phe	His	Gly	Val	Phe	Cys		
1				5					10					15			
Leu	Tyr	Gln	Cys	Arg	Gln	Cys	Leu	Ala	Tyr	His	Val	Cys	Asp	Gly	Gly		
		20						25					30				
Ala	Glu	Cys	Val	Leu	Leu	His	Thr	Pro	Glu	Ser	Val	Ile	Cys	Glu	Leu		
		35					40						45				

[illegible]

ATG	GCT	AGC	CGG	AGG	CGC	AAA	CTT	CGG	AAT	TTC	CTA	AAC	AAG	GAA	TGC	48
Met	Ala	Ser	Arg	Arg	Arg	Lys	Leu	Arg	Asn	Phe	Leu	Asn	Lys	Glu	Cys	
1				5					10					15		
ATA	TGG	ACT	GTT	AAC	CCA	ATG	TCA	GGG	GAC	CAT	ATC	AAG	GTC	TTT	AAC	96
Ile	Trp	Thr	Val	Asn	Pro	Met	Ser	Gly	Asp	His	Ile	Lys	Val	Phe	Asn	
			20					25					30			
GCC	TGC	ACC	TCT	ATC	TCG	CCG	GTG	TAT	GAC	CCT	GAG	CTG	GTA	ACC	AGC	144
Ala	Cys	Thr	Ser	Ile	Ser	Pro	Val	Tyr	Asp	Pro	Glu	Leu	Val	Thr	Ser	
		35					40					45				
TAC	GCA	CTG	AGC	GTG	CCT	GCT	TAC	AAT	GTG	TCT	GTG	GCT	ATC	TTG	CTG	192
Tyr	Ala	Leu	Ser	Val	Pro	Ala	Tyr	Asn	Val	Ser	Val	Ala	Ile	Leu	Leu	
	50					55					60					

204

CAT	AAA	GTC	ATG	GGA	CCG	TGT	GTG	GCT	GTG	GGA	ATT	AAC	GGA	GAA	ATG	240
His	Lys	Val	Met	Gly	Pro	Cys	Val	Ala	Val	Gly	Ile	Asn	Gly	Glu	Met	
65					70					75					80	
ATC	ATG	TAC	GTC	GTA	AGC	CAG	TGT	GTT	TCT	GTG	CGG	CCC	GTC	CCG	GGG	288
Ile	Met	Tyr	Val	Val	Ser	Gln	Cys	Val	Ser	Val	Arg	Pro	Val	Pro	Gly	
				85					90					95		
CGC	GAT	GGT	ATG	GCG	CTC	ATC	TAC	TTT	GGA	CAG	TTT	CTG	GAG	GAA	GCA	336
Arg	Asp	Gly	Met	Ala	Leu	Ile	Tyr	Phe	Gly	Gln	Phe	Leu	Glu	Glu	Ala	
			100					105					110			
TCC	GGA	CTG	AGA	TTT	CCC	TAC	ATT	GCT	CCG	CCG	CCG	TCG	CGC	GAA	CAC	384
Ser	Gly	Leu	Arg	Phe	Pro	Tyr	Ile	Ala	Pro	Pro	Pro	Ser	Arg	Glu	His	
		115					120					125				
GTA	CCT	GAC	CTG	ACC	AGA	CAA	GAA	TTA	GTT	CAT	ACC	TCC	CAG	GTG	GTG	432
Val	Pro	Asp	Leu	Thr	Arg	Gln	Glu	Leu	Val	His	Thr	Ser	Gln	Val	Val	
	130					135					140					
CGC	CGC	GGC	GAC	CTG	ACC	AAT	TGC	ACT	ATG	GGT	CTC	GAA	TTC	AGG	AAT	480
Arg	Arg	Gly	Asp	Leu	Thr	Asn	Cys	Thr	Met	Gly	Leu	Glu	Phe	Arg	Asn	
145					150					155					160	
GTG	AAC	CCT	TTT	GTT	TGG	CTC	GGG	GGC	GGA	TCG	GTG	TGG	CTG	CTG	TTC	528
Val	Asn	Pro	Phe	Val	Trp	Leu	Gly	Gly	Gly	Ser	Val	Trp	Leu	Leu	Phe	
				165					170					175		
TTG	GGC	GTG	GAC	TAC	ATG	GCG	TTC	TGT	CCG	GGT	GTC	GAC	GGA	ATG	CCG	576
Leu	Gly	Val	Asp	Tyr	Met	Ala	Phe	Cys	Pro	Gly	Val	Asp	Gly	Met	Pro	
			180					185					190			
TCG	TTG	GCA	AGA	GTG	GCC	GCC	CTG	CTT	ACC	AGG	TGC	GAC	CAC	CCA	GAC	624
Ser	Leu	Ala	Arg	Val	Ala	Ala	Leu	Leu	Thr	Arg	Cys	Asp	His	Pro	Asp	
		195					200					205				
TGT	GTC	CAC	TGC	CAT	GGA	CTC	CGT	GGA	CAC	GTT	AAT	GTA	TTT	CGT	GGG	672
Cys	Val	His	Cys	His	Gly	Leu	Arg	Gly	His	Val	Asn	Val	Phe	Arg	Gly	
	210					215					220					
TAC	TGT	TCT	GCG	CAG	TCG	CCG	GGT	CTA	TCT	AAC	ATC	TGT	CCC	TGT	ATC	720
Tyr	Cys	Ser	Ala	Gln	Ser	Pro	Gly	Leu	Ser	Asn	Ile	Cys	Pro	Cys	Ile	
225					230					235					240	
AAA	TCA	TGT	GGG	ACC	GGG	AAT	GGA	GTG	ACT	AGG	GTC	ACT	GGA	AAC	AGA	768
Lys	Ser	Cys	Gly	Thr	Gly	Asn	Gly	Val	Thr	Arg	Val	Thr	Gly	Asn	Arg	
				245					250					255		
AAT	TTT	CTG	GGT	CTT	CTG	TTC	GAT	CCC	ATT	GTC	CAG	AGC	AGG	GTA	ACA	816
Asn	Phe	Leu	Gly	Leu	Leu	Phe	Asp	Pro	Ile	Val	Gln	Ser	Arg	Val	Thr	
			260					265					270			
GCT	CTG	AAG	ATA	ACT	AGC	CAC	CCA	ACC	CCC	ACG	CAC	GTC	GAG	AAT	GTG	864
Ala	Leu	Lys	Ile	Thr	Ser	His	Pro	Thr	Pro	Thr	His	Val	Glu	Asn	Val	
			275				280					285				
CTA	ACA	GGA	GTG	CTC	GAC	GAC	GGC	ACC	TTG	GTG	CCG	TCC	GTC	CAA	GGC	912
Leu	Thr	Gly	Val	Leu	Asp	Asp	Gly	Thr	Leu	Val	Pro	Ser	Val	Gln	Gly	
			290			295					300					
ACC	CTG	GGT	CCT	CTT	ACG	AAT	GTC	TGA								939
Thr	Leu	Gly	Pro	Leu	Thr	Asn	Val									
305					310											

205

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Ala Ser Arg Arg Arg Lys Leu Arg Asn Phe Leu Asn Lys Glu Cys
 1 5 10 15
 Ile Trp Thr Val Asn Pro Met Ser Gly Asp His Ile Lys Val Phe Asn
 20 25 30
 Ala Cys Thr Ser Ile Ser Pro Val Tyr Asp Pro Glu Leu Val Thr Ser
 35 40 45
 Tyr Ala Leu Ser Val Pro Ala Tyr Asn Val Ser Val Ala Ile Leu Leu
 50 55 60
 His Lys Val Met Gly Pro Cys Val Ala Val Gly Ile Asn Gly Glu Met
 65 70 75 80
 Ile Met Tyr Val Val Ser Gln Cys Val Ser Val Arg Pro Val Pro Gly
 85 90 95
 Arg Asp Gly Met Ala Leu Ile Tyr Phe Gly Gln Phe Leu Glu Glu Ala
 100 105 110
 Ser Gly Leu Arg Phe Pro Tyr Ile Ala Pro Pro Pro Ser Arg Glu His
 115 120 125
 Val Pro Asp Leu Thr Arg Gln Glu Leu Val His Thr Ser Gln Val Val
 130 135 140
 Arg Arg Gly Asp Leu Thr Asn Cys Thr Met Gly Leu Glu Phe Arg Asn
 145 150 155 160
 Val Asn Pro Phe Val Trp Leu Gly Gly Gly Ser Val Trp Leu Leu Phe
 165 170 175
 Leu Gly Val Asp Tyr Met Ala Phe Cys Pro Gly Val Asp Gly Met Pro
 180 185 190
 Ser Leu Ala Arg Val Ala Ala Leu Leu Thr Arg Cys Asp His Pro Asp
 195 200 205
 Cys Val His Cys His Gly Leu Arg Gly His Val Asn Val Phe Arg Gly
 210 215 220
 Tyr Cys Ser Ala Gln Ser Pro Gly Leu Ser Asn Ile Cys Pro Cys Ile
 225 230 235 240
 Lys Ser Cys Gly Thr Gly Asn Gly Val Thr Arg Val Thr Gly Asn Arg
 245 250 255
 Asn Phe Leu Gly Leu Leu Phe Asp Pro Ile Val Gln Ser Arg Val Thr
 260 265 270
 Ala Leu Lys Ile Thr Ser His Pro Thr Pro Thr His Val Glu Asn Val
 275 280 285
 Leu Thr Gly Val Leu Asp Asp Gly Thr Leu Val Pro Ser Val Gln Gly
 290 295 300
 Thr Leu Gly Pro Leu Thr Asn Val

006019.02900

310

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..86
(D) OTHER INFORMATION:

ATG	GAC	TCA	ACC	AAC	TCT	AAA	AGA	GAG	TTT	ATT	AAG	TCG	GCT	CTG	GAG	48
Met	Asp	Ser	Thr	Asn	Ser	Lys	Arg	Glu	Phe	Ile	Lys	Ser	Ala	Leu	Glu	
1				5				10						15		
GCC	AAC	ATC	AAC	AGG	AGG	GCA	GCT	GTA	TCG	CTA	TTT	GA				86
Ala	Asn	Ile	Asn	Arg	Arg	Ala	Ala	Val	Ser	Leu	Phe					
			20					25								

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Asp Ser Thr Asn Ser Lys Arg Glu Phe Ile Lys Ser Ala Leu Glu
1 5 10 15
Ala Asn Ile Asn Arg Arg Ala Ala Val Ser Leu Phe
20 25

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1743 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: N

207

(iv) ANTI-SENSE: N

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1743

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATG GCA GAA GGC GGT TTT GGA GCG GAC TCG GTG GGG CGC GGC GGA GAA	48
Met Ala Glu Gly Gly Phe Gly Ala Asp Ser Val Gly Arg Gly Gly Glu	
1 5 10 15	
AAG GCC TCT GTG ACT AGG GGA GGC AGG TGG GAC TTG GGG AGC TCG GAC	96
Lys Ala Ser Val Thr Arg Gly Gly Arg Trp Asp Leu Gly Ser Ser Asp	
20 25 30	
GAC GAA TCA AGC ACC TCC ACA ACC AGC ACG GAT ATG GAC GAC CTC CCT	144
Asp Glu Ser Ser Thr Ser Thr Thr Ser Thr Asp Met Asp Asp Leu Pro	
35 40 45	
GAG GAG AGG AAA CCA CTA ACG GGA AAG TCT GTA AAA ACC TCG TAC ATA	192
Glu Glu Arg Lys Pro Leu Thr Gly Lys Ser Val Lys Thr Ser Tyr Ile	
50 55 60	
TAC GAC GTG CCC ACC GTC CCG ACC AGC AAG CCG TGG CAT TTA ATG CAC	240
Tyr Asp Val Pro Thr Val Pro Thr Ser Lys Pro Trp His Leu Met His	
65 70 75 80	
GAC AAC TCC CTC TAC GCA ACG CCT AGG TTT CCG CCC AGA CCT CTC ATA	288
Asp Asn Ser Leu Tyr Ala Thr Pro Arg Phe Pro Pro Arg Pro Leu Ile	
85 90 95	
CGG CAC CCT TCC GAA AAA GGC AGC ATT TTT GCC AGT CGG TTG TCA GCG	336
Arg His Pro Ser Glu Lys Gly Ser Ile Phe Ala Ser Arg Leu Ser Ala	
100 105 110	
ACT GAC GAC GAC TCG GGA GAC TAC GCG CCA ATG GAT CGC TTC GCC TTC	384
Thr Asp Asp Asp Ser Gly Asp Tyr Ala Pro Met Asp Arg Phe Ala Phe	
115 120 125	
CAG AGC CCC AGG GTG TGT GGT CGC CCT CCC CTT CCG CCT CCA AAT CAC	432
Gln Ser Pro Arg Val Cys Gly Arg Pro Pro Leu Pro Pro Pro Asn His	
130 135 140	
CCA CCT CCG GCA ACT AGG CCG GCA GAC GCG TCA ATG GGG GAC GTG GGC	480
Pro Pro Pro Ala Thr Arg Pro Ala Asp Ala Ser Met Gly Asp Val Gly	
145 150 155 160	
TGG GCG GAT CTG CAG GGA CTC AAG AGG ACC CCA AAG GGA TTT TTA AAA	528
Trp Ala Asp Leu Gln Gly Leu Lys Arg Thr Pro Lys Gly Phe Leu Lys	
165 170 175	
ACA TCT ACC AAG GGG GGC AGT CTC AAA GCC CGT GGA CGC GAT GTA GGT	576
Thr Ser Thr Lys Gly Gly Ser Leu Lys Ala Arg Gly Arg Asp Val Gly	
180 185 190	
GAC CGT CTC AGG GAC GGC GGC TTT GCC TTT AGT CCT AGG GGC GTG AAA	624
Asp Arg Leu Arg Asp Gly Gly Phe Ala Phe Ser Pro Arg Gly Val Lys	
195 200 205	
TCT GCC ATA GGG CAA AAC ATT AAA TCA TGG TTG GGG ATC GGA GAA TCA	672
Ser Ala Ile Gly Gln Asn Ile Lys Ser Trp Leu Gly Ile Gly Glu Ser	
210 215 220	
TCG GCG ACT GCT GTC CCC GTC ACC ACG CAG CTT ATG GTA CCG GTG CAC	720

006699-6249560

208

Ser 225	Ala	Thr	Ala	Val 230	Pro	Val	Thr	Thr	Gln 235	Leu	Met	Val	Pro	Val	His 240	
CTC	ATT	AGA	ACG	CCT	GTG	ACC	GTG	GAC	TAC	AGG	AAT	GTT	TAT	TTG	CTT	768
Leu	Ile	Arg	Thr	Pro 245	Val	Thr	Val	Asp	Tyr 250	Arg	Asn	Val	Tyr	Leu 255	Leu	
TAC	TTA	GAG	GGG	GTA	ATG	GGT	GTG	GGC	AAA	TCA	ACG	CTG	GTC	AAC	GCC	816
Tyr	Leu	Glu	Gly 260	Val	Met	Gly	Val	Gly 265	Lys	Ser	Thr	Leu	Val 270	Asn	Ala	
GTG	TGC	GGG	ATC	TTG	CCC	CAG	GAG	AGA	GTG	ACA	AGT	TTT	CCC	GAG	CCC	864
Val	Cys	Gly 275	Ile	Leu	Pro	Gln	Glu	Arg 280	Val	Thr	Ser	Phe 285	Pro	Glu	Pro	
ATG	GTG	TAC	TGG	ACG	AGG	GCA	TTT	ACA	GAT	TGT	TAC	AAG	GAA	ATT	TCC	912
Met	Val 290	Tyr	Trp	Thr	Arg	Ala 295	Phe	Thr	Asp	Cys	Tyr 300	Lys	Glu	Ile	Ser	
CAC	CTG	ATG	AAG	TCT	GGT	AAG	GCG	GGA	GAC	CCG	CTG	ACG	TCT	GCC	AAA	960
His	Leu	Met	Lys	Ser 310	Gly	Lys	Ala	Gly	Asp	Pro 315	Leu	Thr	Ser	Ala	Lys 320	
ATA	TAC	TCA	TGC	CAA	AAC	AAG	TTT	TCG	CTC	CCC	TTC	CGG	ACG	AAC	GCC	1008
Ile	Tyr	Ser	Cys 325	Gln	Asn	Lys	Phe	Ser 330	Leu	Pro	Phe	Arg	Thr	Asn 335	Ala	
ACC	GCT	ATC	CTG	CGA	ATG	ATG	CAG	CCC	TGG	AAC	GTT	GGG	GGT	GGG	TCT	1056
Thr	Ala	Ile	Leu 340	Arg	Met	Met	Gln	Pro 345	Trp	Asn	Val	Gly 350	Gly	Gly	Ser	
GGG	AGG	GGC	ACT	CAC	TGG	TGC	GTC	TTT	GAT	AGG	CAT	CTC	CTC	TCC	CCA	1104
Gly	Arg	Gly 355	Thr	His	Trp	Cys	Val 360	Phe	Asp	Arg	His 365	Leu	Leu	Ser	Pro	
GCA	GTG	GTG	TTC	CCT	CTC	ATG	CAC	CTG	AAG	CAC	GGC	CGC	CTA	TCT	TTT	1152
Ala	Val 370	Val	Phe	Pro	Leu	Met 375	His	Leu	Lys	His 380	Gly	Arg	Leu	Ser	Phe	
GAT	CAC	TTC	TTT	CAA	TTA	CTT	TCC	ATC	TTT	AGA	GCC	ACA	GAA	GGC	GAC	1200
Asp	His	Phe	Phe 385	Gln	Leu 390	Leu	Ser	Ile	Phe 395	Arg	Ala	Thr	Glu	Gly 400	Asp	
GTG	GTC	GCC	ATT	CTC	ACC	CTC	TCC	AGC	GCC	GAG	TCG	TTG	CGG	CGG	GTC	1248
Val	Val	Ala	Ile 405	Leu	Thr	Leu	Ser	Ser 410	Ala	Glu	Ser	Leu	Arg	Arg 415	Val	
AGG	GCG	AGG	GGA	AGA	AAG	AAC	GAC	GGG	ACG	GTG	GAG	CAA	AAC	TAC	ATC	1296
Arg	Ala	Arg	Gly 420	Arg	Lys	Asn	Asp	Gly 425	Thr	Val	Glu	Gln 430	Asn	Tyr	Ile	
AGA	GAA	TTG	GCG	TGG	GCT	TAT	CAC	GCC	GTG	TAC	TGT	TCA	TGG	ATC	ATG	1344
Arg	Glu 435	Leu	Ala	Trp	Ala	Tyr 440	His	Ala	Val	Tyr	Cys 445	Ser	Trp	Ile	Met	
TTG	CAG	TAC	ATC	ACT	GTG	GAG	CAG	ATG	GTA	CAA	CTA	TGC	GTA	CAA	ACC	1392
Leu	Gln 450	Tyr	Ile	Thr	Val 455	Gln	Met	Val	Gln 460	Leu	Cys	Val	Gln	Thr		
ACA	AAT	ATT	CCG	GAA	ATC	TGC	TTC	CGC	AGC	GTG	CGC	CTG	GCA	CAC	AAG	1440
Thr	Asn	Ile	Pro 465	Glu	Ile 470	Cys	Phe	Arg	Ser 475	Val	Arg	Leu	Ala	His 480	Lys	
GAG	GAA	ACT	TTG	AAA	AAC	CTT	CAC	GAG	CAG	AGC	ATG	CTA	CCT	ATG	ATC	1488
Glu	Glu	Thr	Leu 485	Lys	Asn	Leu	His	Glu 490	Gln	Ser	Met	Leu	Pro	Met 495	Ile	

006250" 6249960

209

ACC GGT GTA CTG GAT CCC GTG AGA CAT CAT CCC GTC GTG ATC GAG CTT 1536
 Thr Gly Val Leu Asp Pro Val Arg His His Pro Val Val Ile Glu Leu
 500 505 510

TGC TTT TGT TTC TTC ACA GAG CTG AGA AAA TTA CAA TTT ATC GTA GCC 1564
 Cys Phe Cys Phe Phe Thr Glu Leu Arg Lys Leu Gln Phe Ile Val Ala
 515 520 525

GAC GCG GAT AAG TTC CAC GAC GAC GTA TGC GGC CTG TGG ACC GAA ATC 1632
 Asp Ala Asp Lys Phe His Asp Asp Val Cys Gly Leu Trp Thr Glu Ile
 530 535 540

TAC AGG CAG ATC CTG TCC AAT CCG GCT ATT AAA CCC AGG GCC ATC AAC 1680
 Tyr Arg Gln Ile Leu Ser Asn Pro Ala Ile Lys Pro Arg Ala Ile Asn
 545 550 555 560

TGG CCA GCA TTA GAG AGC CAG TCT AAA GCA GTT AAT CAC CTA GAG GAG 1728
 Trp Pro Ala Leu Glu Ser Gln Ser Lys Ala Val Asn His Leu Glu Glu
 565 570 575

ACA TGC AGG GTC TAG 1743
 Thr Cys Arg Val
 580

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 580 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Ala Glu Gly Gly Phe Gly Ala Asp Ser Val Gly Arg Gly Gly Glu
 1 5 10 15

Lys Ala Ser Val Thr Arg Gly Gly Arg Trp Asp Leu Gly Ser Ser Asp
 20 25 30

Asp Glu Ser Ser Thr Ser Thr Thr Ser Thr Asp Met Asp Asp Leu Pro
 35 40 45

Glu Glu Arg Lys Pro Leu Thr Gly Lys Ser Val Lys Thr Ser Tyr Ile
 50 55 60

Tyr Asp Val Pro Thr Val Pro Thr Ser Lys Pro Trp His Leu Met His
 65 70 75 80

Asp Asn Ser Leu Tyr Ala Thr Pro Arg Phe Pro Pro Arg Pro Leu Ile
 85 90 95

Arg His Pro Ser Glu Lys Gly Ser Ile Phe Ala Ser Arg Leu Ser Ala
 100 105 110

Thr Asp Asp Asp Ser Gly Asp Tyr Ala Pro Met Asp Arg Phe Ala Phe
 115 120 125

Gln Ser Pro Arg Val Cys Gly Arg Pro Pro Leu Pro Pro Pro Asn His
 130 135 140

Pro Pro Pro Ala Thr Arg Pro Ala Asp Ala Ser Met Gly Asp Val Gly
 145 150 155 160

Trp Ala Asp Leu Gln Gly Leu Lys Arg Thr Pro Lys Gly Phe Leu Lys

006250 625060

165								170				175			
Thr	Ser	Thr	Lys 180	Gly	Gly	Ser	Leu	Lys 185	Ala	Arg	Gly	Arg	Asp 190	Val	Gly
Asp	Arg	Leu 195	Arg	Asp	Gly	Gly	Phe 200	Ala	Phe	Ser	Pro	Arg 205	Gly	Val	Lys
Ser	Ala 210	Ile	Gly	Gln	Asn	Ile 215	Lys	Ser	Trp	Leu	Gly 220	Ile	Gly	Glu	Ser
Ser 225	Ala	Thr	Ala	Val	Pro 230	Val	Thr	Thr	Gln	Leu 235	Met	Val	Pro	Val	His 240
Leu	Ile	Arg	Thr	Pro 245	Val	Thr	Val	Asp	Tyr 250	Arg	Asn	Val	Tyr	Leu 255	Leu
Tyr	Leu	Glu	Gly 260	Val	Met	Gly	Val	Gly 265	Lys	Ser	Thr	Leu	Val	Asn	Ala
Val	Cys	Gly 275	Ile	Leu	Pro	Gln	Glu 280	Arg	Val	Thr	Ser	Phe 285	Pro	Glu	Pro
Met	Val 290	Tyr	Trp	Thr	Arg	Ala 295	Phe	Thr	Asp	Cys	Tyr 300	Lys	Glu	Ile	Ser
His 305	Leu	Met	Lys	Ser	Gly 310	Lys	Ala	Gly	Asp	Pro 315	Leu	Thr	Ser	Ala	Lys 320
Ile	Tyr	Ser	Cys	Gln 325	Asn	Lys	Phe	Ser	Leu 330	Pro	Phe	Arg	Thr	Asn 335	Ala
Thr	Ala	Ile	Leu 340	Arg	Met	Met	Gln	Pro 345	Trp	Asn	Val	Gly	Gly 350	Gly	Ser
Gly	Arg	Gly 355	Thr	His	Trp	Cys	Val 360	Phe	Asp	Arg	His	Leu 365	Leu	Ser	Pro
Ala	Val 370	Val	Phe	Pro	Leu	Met 375	His	Leu	Lys	His	Gly 380	Arg	Leu	Ser	Phe
Asp 385	His	Phe	Phe	Gln 390	Leu	Leu	Ser	Ile	Phe	Arg 395	Ala	Thr	Glu	Gly	Asp 400
Val	Val	Ala	Ile	Leu 405	Thr	Leu	Ser	Ser	Ala 410	Glu	Ser	Leu	Arg	Arg 415	Val
Arg	Ala	Arg	Gly 420	Arg	Lys	Asn	Asp	Gly 425	Thr	Val	Glu	Gln	Asn 430	Tyr	Ile
Arg	Glu	Leu 435	Ala	Trp	Ala	Tyr	His 440	Ala	Val	Tyr	Cys	Ser 445	Trp	Ile	Met
Leu	Gln 450	Tyr	Ile	Thr	Val	Glu 455	Gln	Met	Val	Gln	Leu 460	Cys	Val	Gln	Thr
Thr 465	Asn	Ile	Pro	Glu	Ile 470	Cys	Phe	Arg	Ser	Val 475	Arg	Leu	Ala	His	Lys 480
Glu	Glu	Thr	Leu	Lys 485	Asn	Leu	His	Glu	Gln 490	Ser	Met	Leu	Pro	Met 495	Ile
Thr	Gly	Val	Leu 500	Asp	Pro	Val	Arg	His 505	His	Pro	Val	Val	Ile 510	Glu	Leu
Cys	Phe	Cys 515	Phe	Phe	Thr	Glu	Leu 520	Arg	Lys	Leu	Gln	Phe 525	Ile	Val	Ala

WO 96/06159

211

Asp Ala Asp Lys Phe His Asp Asp Val Cys Gly Leu Trp Thr Glu Ile
 530 535 540

Tyr Arg Gln Ile Leu Ser Asn Pro Ala Ile Lys Pro Arg Ala Ile Asn
 545 550 555 560

Trp Pro Ala Leu Glu Ser Gln Ser Lys Ala Val Asn His Leu Glu Glu
 565 570 575

Thr Cys Arg Val
 580

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2193 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..2193
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATG CAG GGT CTA GCC TTC TTG GCG GCC CTT GCA TGC TGG CGA TGC ATA	48
Met Gln Gly Leu Ala Phe Leu Ala Ala Leu Ala Cys Trp Arg Cys Ile	15
1 5 10	
TCG TTG ACA TGT GGA GCC ACT GGC GCG TTG CCG ACA ACG GCG ACG ACA	96
Ser Leu Thr Cys Gly Ala Thr Gly Ala Leu Pro Thr Thr Ala Thr Thr	30
20 25	
ATA ACC CGC TCC GCC ACG CAG CTC ATC AAT GGG AGA ACC AAC CTC TCC	144
Ile Thr Arg Ser Ala Thr Gln Leu Ile Asn Gly Arg Thr Asn Leu Ser	45
35 40	
ATA GAA CTG GAA TTC AAC GGC ACT AGT TTT TTT CTA AAT TGG CAA AAT	192
Ile Glu Leu Glu Phe Asn Gly Thr Ser Phe Phe Leu Asn Trp Gln Asn	60
50 55	
CTG TTG AAT GTG ATC ACG GAG CCG GCC CTG ACA GAG TTG TGG ACC TCC	240
Leu Leu Asn Val Ile Thr Glu Pro Ala Leu Thr Glu Leu Trp Thr Ser	80
65 70	
GCC GAA GTC GCC GAG GAC CTC AGG GTA ACT CTG AAA AAG AGG CAA AGT	288
Ala Glu Val Ala Glu Asp Leu Arg Val Thr Leu Lys Lys Arg Gln Ser	95
85 90	
CTT TTT TTC CCC AAC AAG ACA GTT GTG ATC TCT GGA GAC GGC CAT CGC	336
Leu Phe Phe Pro Asn Lys Thr Val Val Ile Ser Gly Asp Gly His Arg	110
100 105	
TAT ACG TGC GAG GTG CCG ACG TCG TCG CAA ACT TAT AAC ATC ACC AAG	384
Tyr Thr Cys Glu Val Pro Thr Ser Ser Gln Thr Tyr Asn Ile Thr Lys	125
115 120	
GGC TTT AAC TAT AGC GCT CTG CCC GGG CAC CTT GGC GGA TTT GGG ATC	432

006290-52120950

Gly 130	Phe 130	Asn	Tyr	Ser	Ala	Leu 135	Pro	Gly	His	Leu	Gly 140	Gly	Phe	Gly	Ile	
AAC Asn 145	GCG Ala	CGT Arg	CTG Leu	GTA Val	CTG Leu 150	GGT Gly	GAT Asp	ATC Ile	TTC Phe	GCA Ala 155	TCA Ser	AAA Lys	TGG Trp	TCG Ser	CTA Leu 160	480
TTC Phe	GCG Ala	AGG Arg	GAC Asp	ACC Thr 165	CCA Pro	GAG Glu	TAT Tyr	CGG Arg	GTG Val 170	TTT Phe	TAC Tyr	CCA Pro	ATG Met	AAT Asn 175	GTC Val	528
ATG Met	GCC Ala	GTC Val	AAG Lys 180	TTT Phe	TCC Ser	ATA Ile	TCC Ser	ATT Ile 185	GGC Gly	AAC Asn	AAC Asn	GAG Glu 190	TCC Ser	GGC Gly	GTA Val	576
GCG Ala	CTC Leu	TAT Tyr 195	GGA Gly	GTG Val	GTG Val	TCG Ser	GAA Glu 200	GAT Asp	TTC Phe	GTG Val	GTC Val	GTC Val 205	ACG Thr	CTC Leu	CAC His	624
AAC Asn 210	AGG Arg	TCC Ser	AAA Lys	GAG Glu	GCT Ala	AAC Asn 215	GAG Glu	ACG Thr	GCG Ala	TCC Ser	CAT His 220	CTT Leu	CTG Leu	TTC Phe	GGT Gly	672
CTC Leu 225	CCG Pro	GAT Asp	TCA Ser	CTG Leu	CCA Pro 230	TCT Ser	CTG Leu	AAG Lys	GGC Gly	CAT His 235	GCC Ala	ACC Thr	TAT Tyr	GAT Asp	GAA Glu 240	720
CTC Leu	ACG Thr	TTC Phe	GCC Ala	CGA Arg 245	AAC Asn	GCA Ala	AAA Lys	TAT Tyr	GCG Ala 250	CTA Leu	GTG Val	GCG Ala	ATC Ile	CTG Leu 255	CCT Pro	768
AAA Lys	GAT Asp	TCT Ser	TAC Tyr 260	CAG Gln	ACA Thr	CTC Leu	CTT Leu	ACA Thr 265	GAG Glu	AAT Asn	TAC Tyr	ACT Thr	CGC Arg 270	ATA Ile	TTT Phe	816
CTG Leu	AAC Asn	ATG Met 275	ACG Thr	GAG Glu	TCG Ser	ACG Thr	CCC Pro 280	CTC Leu	GAG Glu	TTC Phe	ACG Thr	CGG Arg 285	ACG Thr	ATC Ile	CAG Gln	864
ACC Thr 290	AGG Arg	ATC Ile	GTA Val	TCA Ser	ATC Ile	GAG Glu 295	GCC Ala	AGG Arg	CGC Arg	GCC Ala	TGC Cys 300	GCA Ala	GCT Ala	CAA Gln	GAG Glu	912
GCG Ala 305	GCG Ala	CCG Pro	GAC Asp	ATA Ile	TTC Phe 310	TTG Leu	GTG Val	TTG Leu	TTT Phe	CAG Gln 315	ATG Met	TTG Leu	GTG Val	GCA Ala 320	CAC His	960
TTT Phe	CTT Leu	GTT Val	GCG Ala	CGG Arg 325	GGC Gly	ATT Ile	GCC Ala	GAG Glu	CAC His 330	CGA Arg	TTT Phe	GTG Val	GAG Glu	GTG Val 335	GAC Asp	1008
TGC Cys	GTG Val	TGT Cys	CGG Arg 340	CAG Gln	TAT Tyr	GCG Ala	GAA Glu	CTG Leu 345	TAT Tyr	TTT Phe	CTC Leu	CGC Arg 350	CGC Arg	ATC Ile	TCG Ser	1056
CGT Arg	CTG Leu	TGC Cys 355	ATG Met	CCC Pro	ACG Thr	TTC Phe	ACC Thr 360	ACT Thr	GTC Val	GGG Gly	TAT Tyr	AAC Asn 365	CAC His	ACC Thr	ACC Thr	1104
CTT Leu	GGC Gly 370	GCT Ala	GTG Val	GCC Ala	GCC Ala	ACA Thr 375	CAA Gln	ATA Ile	GCT Ala	CGC Arg	GTG Val 380	TCC Ser	GCC Ala	ACG Thr	AAG Lys	1152
TTG Leu 385	GCC Ala	AGT Ser	TTG Leu	CCC Pro	CGC Arg 390	TCT Ser	TCC Ser	CAG Gln	GAA Glu 395	ACA Thr	GTG Val	CTG Leu	GCC Ala	ATG Met	GTC Val 400	1200

CAG	CTT	GGC	GCC	CGT	GAT	GGC	GCC	GTC	CCT	TCC	TCC	ATT	CTG	GAG	GGC	1248
Gln	Leu	Gly	Ala	Arg	Asp	Gly	Ala	Val	Pro	Ser	Ser	Ile	Leu	Glu	Gly	
				405					410					415		
ATT	GCT	ATG	GTC	GTC	GAA	CAT	ATG	TAT	ACC	GCC	TAC	ACT	TAT	GTG	TAC	1296
Ile	Ala	Met	Val	Val	Glu	His	Met	Tyr	Thr	Ala	Tyr	Thr	Tyr	Val	Tyr	
			420					425					430			
ACA	CTC	GGC	GAT	ACT	GAA	AGA	AAA	TTA	ATG	TTG	GAC	ATA	CAC	ACG	GTC	1344
Thr	Leu	Gly	Asp	Thr	Glu	Arg	Lys	Leu	Met	Leu	Asp	Ile	His	Thr	Val	
		435					440					445				
CTC	ACC	GAC	AGC	TGC	CCG	CCC	AAA	GAC	TCC	GGA	GTA	TCA	GAA	AAG	CTA	1392
Leu	Thr	Asp	Ser	Cys	Pro	Pro	Lys	Asp	Ser	Gly	Val	Ser	Glu	Lys	Leu	
		450				455					460					
CTG	AGA	ACA	TAT	TTG	ATG	TTC	ACA	TCA	ATG	TGT	ACC	AAC	ATA	GAG	CTG	1440
Leu	Arg	Thr	Tyr	Leu	Met	Phe	Thr	Ser	Met	Cys	Thr	Asn	Ile	Glu	Leu	
465					470					475					480	
GGC	GAA	ATG	ATC	GCC	CGC	TTT	TCC	AAA	CCG	GAC	AGC	CTT	AAC	ATC	TAT	1488
Gly	Glu	Met	Ile	Ala	Arg	Phe	Ser	Lys	Pro	Asp	Ser	Leu	Asn	Ile	Tyr	
				485					490					495		
AGG	GCA	TTC	TCC	CCC	TGC	TTT	CTA	GGA	CTA	AGG	TAC	GAT	TTG	CAT	CCA	1536
Arg	Ala	Phe	Ser	Pro	Cys	Phe	Leu	Gly	Leu	Arg	Tyr	Asp	Leu	His	Pro	
			500					505					510			
GCC	AAG	TTG	CGC	GCC	GAG	GCG	CCG	CAG	TCG	TCC	GCT	CTG	ACG	CGG	ACT	1584
Ala	Lys	Leu	Arg	Ala	Glu	Ala	Pro	Gln	Ser	Ser	Ala	Leu	Thr	Arg	Thr	
		515					520					525				
GCC	GTT	GCC	AGA	GGA	ACA	TCG	GGA	TTC	GCA	GAA	TTG	CTC	CAC	GCG	CTG	1632
Ala	Val	Ala	Arg	Gly	Thr	Ser	Gly	Phe	Ala	Glu	Leu	Leu	His	Ala	Leu	
	530					535					540					
CAC	CTC	GAT	AGC	TTA	AAT	TTA	ATT	CCG	GCG	ATT	AAC	TGT	TCA	AAG	ATT	1680
His	Leu	Asp	Ser	Leu	Asn	Leu	Ile	Pro	Ala	Ile	Asn	Cys	Ser	Lys	Ile	
545					550					555					560	
ACA	GCC	GAC	AAG	ATA	ATA	GCT	ACG	GTA	CCC	TTG	CCT	CAC	GTC	ACG	TAT	1728
Thr	Ala	Asp	Lys	Ile	Ile	Ala	Thr	Val	Pro	Leu	Pro	His	Val	Thr	Tyr	
				565					570					575		
ATC	ATC	AGT	TCC	GAA	GCA	CTC	TCG	AAC	GCT	GTT	GTC	TAC	GAG	GTG	TCG	1776
Ile	Ile	Ser	Ser	Glu	Ala	Leu	Ser	Asn	Ala	Val	Val	Tyr	Glu	Val	Ser	
			580					585					590			
GAG	ATC	TTC	CTC	AAG	AGT	GCC	ATG	TTT	ATA	TCT	GCT	ATC	AAA	CCC	GAT	1824
Glu	Ile	Phe	Leu	Lys	Ser	Ala	Met	Phe	Ile	Ser	Ala	Ile	Lys	Pro	Asp	
		595					600					605				
TGC	TCC	GGC	TTT	AAC	TTT	TCT	CAG	ATT	GAT	AGG	CAC	ATT	CCC	ATA	GTC	1872
Cys	Ser	Gly	Phe	Asn	Phe	Ser	Gln	Ile	Asp	Arg	His	Ile	Pro	Ile	Val	
	610					615					620					
TAC	AAC	ATC	AGC	ACA	CCA	AGA	AGA	GGT	TGC	CCC	CTT	TGT	GAC	TCT	GTA	1920
Tyr	Asn	Ile	Ser	Thr	Pro	Arg	Arg	Gly	Cys	Pro	Leu	Cys	Asp	Ser	Val	
625					630				635						640	
ATC	ATG	AGC	TAC	GAT	GAG	AGC	GAT	GGC	CTG	CAG	TCT	CTC	ATG	TAT	GTC	1968
Ile	Met	Ser	Tyr	Asp	Glu	Ser	Asp	Gly	Leu	Gln	Ser	Leu	Met	Tyr	Val	
				645				650						655		
ACT	AAT	GAA	AGG	GTG	CAG	ACC	AAC	CTC	TTT	TTA	GAT	AAG	TCA	CCT	TTC	2016
Thr	Asn	Glu	Arg	Val	Gln	Thr	Asn	Leu	Phe	Leu	Asp	Lys	Ser	Pro	Phe	
			660					665					670			

Asn 210	Arg	Ser	Lys	Glu	Ala	Asn 215	Glu	Thr	Ala	Ser	His 220	Leu	Leu	Phe	Gly
Leu 225	Pro	Asp	Ser	Leu	Pro 230	Ser	Leu	Lys	Gly	His 235	Ala	Thr	Tyr	Asp	Glu 240
Leu	Thr	Phe	Ala	Arg 245	Asn	Ala	Lys	Tyr	Ala 250	Leu	Val	Ala	Ile	Leu 255	Pro
Lys	Asp	Ser	Tyr 260	Gln	Thr	Leu	Leu	Thr 265	Glu	Asn	Tyr	Thr	Arg 270	Ile	Phe
Leu	Asn	Met 275	Thr	Glu	Ser	Thr	Pro 280	Leu	Glu	Phe	Thr	Arg 285	Thr	Ile	Gln
Thr	Arg 290	Ile	Val	Ser	Ile	Glu 295	Ala	Arg	Arg	Ala	Cys 300	Ala	Ala	Gln	Glu
Ala 305	Ala	Pro	Asp	Ile	Phe 310	Leu	Val	Leu	Phe	Gln 315	Met	Leu	Val	Ala	His 320
Phe	Leu	Val	Ala	Arg 325	Gly	Ile	Ala	Glu	His 330	Arg	Phe	Val	Glu	Val 335	Asp
Cys	Val	Cys	Arg 340	Gln	Tyr	Ala	Glu	Leu 345	Tyr	Phe	Leu	Arg	Arg 350	Ile	Ser
Arg	Leu	Cys 355	Met	Pro	Thr	Phe	Thr 360	Thr	Val	Gly	Tyr	Asn 365	His	Thr	Thr
Leu	Gly 370	Ala	Val	Ala	Ala	Thr 375	Gln	Ile	Ala	Arg	Val 380	Ser	Ala	Thr	Lys
Leu 385	Ala	Ser	Leu	Pro	Arg 390	Ser	Ser	Gln	Glu	Thr 395	Val	Leu	Ala	Met	Val 400
Gln	Leu	Gly	Ala	Arg 405	Asp	Gly	Ala	Val	Pro 410	Ser	Ser	Ile	Leu	Glu 415	Gly
Ile	Ala	Met 420	Val	Val	Glu	His	Met	Tyr 425	Thr	Ala	Tyr	Thr	Tyr 430	Val	Tyr
Thr	Leu	Gly 435	Asp	Thr	Glu	Arg	Lys 440	Leu	Met	Leu	Asp	Ile 445	His	Thr	Val
Leu	Thr	Asp	Ser	Cys	Pro	Pro 455	Lys	Asp	Ser	Gly	Val 460	Ser	Glu	Lys	Leu
Leu 465	Arg	Thr	Tyr	Leu	Met 470	Phe	Thr	Ser	Met	Cys 475	Thr	Asn	Ile	Glu	Leu 480
Gly	Glu	Met	Ile	Ala	Arg 485	Phe	Ser	Lys	Pro 490	Asp	Ser	Leu	Asn	Ile 495	Tyr
Arg	Ala	Phe	Ser 500	Pro	Cys	Phe	Leu	Gly 505	Leu	Arg	Tyr	Asp	Leu 510	His	Pro
Ala	Lys	Leu 515	Arg	Ala	Glu	Ala	Pro 520	Gln	Ser	Ser	Ala	Leu 525	Thr	Arg	Thr
Ala 530	Val	Ala	Arg	Gly	Thr	Ser 535	Gly	Phe	Ala	Glu	Leu 540	Leu	His	Ala	Leu
His 545	Leu	Asp	Ser	Leu	Asn 550	Leu	Ile	Pro	Ala	Ile 555	Asn	Cys	Ser	Lys	Ile 560
Thr	Ala	Asp	Lys	Ile	Ile	Ala	Thr	Val	Pro	Leu	Pro	His	Val	Thr	Tyr

217

Leu	Phe	Glu	Lys	Phe	Gly	Thr	Pro	Asp	Ser	Ser	Thr	Leu	Pro	Leu	Tyr	
	50					55					60					
GCG	GCT	AGG	CAC	CCG	GAA	CTA	TCG	TTG	CTA	CGG	ATC	ATG	CTC	TCA	CCG	240
Ala	Ala	Arg	His	Pro	Glu	Leu	Ser	Leu	Leu	Arg	Ile	Met	Leu	Ser	Pro	
65					70					75					80	
CAC	CCC	TAC	GCG	TTA	AGA	AGC	CAC	TTG	TGC	GTA	GGC	GAA	GAG	ACC	GCA	288
His	Pro	Tyr	Ala	Leu	Arg	Ser	His	Leu	Cys	Val	Gly	Glu	Glu	Thr	Ala	
				85					90					95		
TCT	CTT	GGC	GTT	TAC	CTG	CAC	TCC	AAG	CCA	GTC	GTA	CGC	GGC	CAC	GAA	336
Ser	Leu	Gly	Val	Tyr	Leu	His	Ser	Lys	Pro	Val	Val	Arg	Gly	His	Glu	
			100					105					110			
TTC	GAG	GAC	ACG	CAG	ATA	CTA	CCG	GAG	TGC	CGG	CTG	GCC	ATA	ACG	AGC	384
Phe	Glu	Asp	Thr	Gln	Ile	Leu	Pro	Glu	Cys	Arg	Leu	Ala	Ile	Thr	Ser	
		115					120					125				
GAC	CAG	TCT	TAT	ACC	AAC	TTT	AAG	ATT	ATA	GAT	CTG	CCA	GCG	GGA	TGC	432
Asp	Gln	Ser	Tyr	Thr	Asn	Phe	Lys	Ile	Ile	Asp	Leu	Pro	Ala	Gly	Cys	
	130					135					140					
CGT	CGC	GTC	CCC	ATA	CAC	GCC	GCG	AAC	AAG	CGT	GTC	GTC	ATC	GAC	GAG	480
Arg	Arg	Val	Pro	Ile	His	Ala	Ala	Asn	Lys	Arg	Val	Val	Ile	Asp	Glu	
145					150					155					160	
GCC	GCC	AAC	CGC	ATA	AAG	GTG	TTT	GAC	CCA	GAG	TCG	CCT	TTA	CCG	CGT	528
Ala	Ala	Asn	Arg	Ile	Lys	Val	Phe	Asp	Pro	Glu	Ser	Pro	Leu	Pro	Arg	
				165					170					175		
CAC	CCC	ATA	ACA	CCC	CGT	GCC	GGT	CAG	ACC	AGA	TCT	ATA	CTG	AAA	CAC	576
His	Pro	Ile	Thr	Pro	Arg	Ala	Gly	Gln	Thr	Arg	Ser	Ile	Leu	Lys	His	
			180					185					190			
AAC	ATC	GCA	CAG	GTT	TGC	GAA	CGG	GAT	ATC	GTG	TCA	CTT	AAC	ACA	GAC	624
Asn	Ile	Ala	Gln	Val	Cys	Glu	Arg	Asp	Ile	Val	Ser	Leu	Asn	Thr	Asp	
		195					200					205				
AAC	GAG	GCC	GCG	TCT	ATG	TTC	TAC	ATG	ATT	GGA	CTC	AGG	CGG	CCG	AGA	672
Asn	Glu	Ala	Ala	Ser	Met	Phe	Tyr	Met	Ile	Gly	Leu	Arg	Arg	Pro	Arg	
	210					215					220					
CTC	GGA	GAA	AGC	CCG	GTC	TGT	GAC	TTC	AAC	ACC	GTT	ACC	ATC	ATG	GAG	720
Leu	Gly	Glu	Ser	Pro	Val	Cys	Asp	Phe	Asn	Thr	Val	Thr	Ile	Met	Glu	
225					230					235					240	
CGT	GCT	AAC	AAC	TCG	ATA	ACT	TTT	CTA	CCC	AAG	CTA	AAA	CTG	AAC	CGG	768
Arg	Ala	Asn	Asn	Ser	Ile	Thr	Phe	Leu	Pro	Lys	Leu	Lys	Leu	Asn	Arg	
				245					250				255			
CTA	CAA	CAC	CTG	TTC	CTG	AAG	CAC	GTG	TTG	CTG	CGC	AGC	ATG	GGG	CTG	816
Leu	Gln	His	Leu	Phe	Leu	Lys	His	Val	Leu	Leu	Arg	Ser	Met	Gly	Leu	
			260					265					270			
GAA	AAC	ATC	GTG	TCG	TGT	TTC	TCA	TCG	CTG	TAC	GGC	GCA	GAA	CTT	GCC	864
Glu	Asn	Ile	Val	Ser	Cys	Phe	Ser	Ser	Leu	Tyr	Gly	Ala	Glu	Leu	Ala	
		275					280					285				
CCT	GCG	AAA	ACA	CAC	GAG	CGG	GAG	TTC	TTC	GGC	GCT	CTG	CTA	GAA	AGA	912
Pro	Ala	Lys	Thr	His	Glu	Arg	Glu	Phe	Phe	Gly	Ala	Leu	Leu	Glu	Arg	
	290					295				300						
CTC	AAA	CGT	CGG	GTG	GAG	GAC	GCG	GTC	TTC	TGC	CTG	AAT	ACC	ATA	GAG	960
Leu	Lys	Arg	Arg	Val	Glu	Asp	Ala	Val	Phe	Cys	Leu	Asn	Thr	Ile	Glu	
305					310					315					320	

006293" 64E4960

218

GAT	TTC	CCG	TTT	AGG	GAA	CCC	ATT	CGC	CAA	CCC	CCA	GAT	TGT	TCC	AAG	1008
Asp	Phe	Pro	Phe	Arg	Glu	Pro	Ile	Arg	Gln	Pro	Pro	Asp	Cys	Ser	Lys	
				325					330					335		
GTG	CTT	ATA	GAA	GCC	ATG	GAA	AAG	TAC	TTT	ATG	ATG	TGT	AGC	CCC	AAA	1056
Val	Leu	Ile	Glu	Ala	Met	Glu	Lys	Tyr	Phe	Met	Met	Cys	Ser	Pro	Lys	
			340					345					350			
GAC	CGT	CAA	AGC	GCC	GCA	TGG	CTA	GGT	GCA	GGG	GTG	GTC	GAA	CTG	ATA	1104
Asp	Arg	Gln	Ser	Ala	Ala	Trp	Leu	Gly	Ala	Gly	Val	Val	Glu	Leu	Ile	
		355					360					365				
TGT	GAC	GGC	AAT	CCA	CTT	TCT	GAG	GTG	CTC	GGA	TTT	CTT	GCC	AAG	TAT	1152
Cys	Asp	Gly	Asn	Pro	Leu	Ser	Glu	Val	Leu	Gly	Phe	Leu	Ala	Lys	Tyr	
	370					375					380					
ATG	CCC	ATA	CAA	AAA	GAA	TGC	ACA	GGA	AAC	CTT	TTA	AAA	ATC	TAC	GCT	1200
Met	Pro	Ile	Gln	Lys	Glu	Cys	Thr	Gly	Asn	Leu	Leu	Lys	Ile	Tyr	Ala	
	385				390				395						400	
TTA	TTG	ACC	GTC	TAA												1215
Leu	Leu	Thr	Val													

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 404 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met	Leu	Arg	Val	Pro	Asp	Val	Lys	Ala	Ser	Leu	Val	Glu	Gly	Ala	Ala	
1				5					10					15		
Arg	Leu	Ser	Thr	Gly	Glu	Arg	Val	Phe	His	Val	Leu	Thr	Ser	Pro	Ala	
		20					25					30				
Val	Ala	Ala	Met	Val	Gly	Val	Ser	Asn	Pro	Glu	Val	Pro	Met	Pro	Leu	
	35					40					45					
Leu	Phe	Glu	Lys	Phe	Gly	Thr	Pro	Asp	Ser	Ser	Thr	Leu	Pro	Leu	Tyr	
	50				55						60					
Ala	Ala	Arg	His	Pro	Glu	Leu	Ser	Leu	Leu	Arg	Ile	Met	Leu	Ser	Pro	
	65				70				75					80		
His	Pro	Tyr	Ala	Leu	Arg	Ser	His	Leu	Cys	Val	Gly	Glu	Glu	Thr	Ala	
			85					90						95		
Ser	Leu	Gly	Val	Tyr	Leu	His	Ser	Lys	Pro	Val	Val	Arg	Gly	His	Glu	
		100					105						110			
Phe	Glu	Asp	Thr	Gln	Ile	Leu	Pro	Glu	Cys	Arg	Leu	Ala	Ile	Thr	Ser	
		115				120					125					
Asp	Gln	Ser	Tyr	Thr	Asn	Phe	Lys	Ile	Ile	Asp	Leu	Pro	Ala	Gly	Cys	
	130					135					140					
Arg	Arg	Val	Pro	Ile	His	Ala	Ala	Asn	Lys	Arg	Val	Val	Ile	Asp	Glu	
	145				150				155					160		
Ala	Ala	Asn	Arg	Ile	Lys	Val	Phe	Asp	Pro	Glu	Ser	Pro	Leu	Pro	Arg	

006220-62720960

220

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATG	GCA	GCG	CTC	GAG	GGC	CCC	CTA	CTA	CTG	CCA	CCG	AGC	GCC	TCC	CTG	48
Met	Ala	Ala	Leu	Glu	Gly	Pro	Leu	Leu	Leu	Pro	Pro	Ser	Ala	Ser	Leu	
1				5					10					15		
ACG	ACG	AGT	CCG	CAG	ACC	ACG	TGT	TAT	CAA	GCG	ACT	TGG	GAA	TCA	CAG	96
Thr	Thr	Ser	Pro	Gln	Thr	Thr	Cys	Tyr	Gln	Ala	Thr	Trp	Glu	Ser	Gln	
			20					25					30			
CTG	GAA	ATA	TTC	TGC	TGT	CTG	GCC	ACC	AAC	TCG	CAC	CTG	CAG	GCA	GAG	144
Leu	Glu	Ile	Phe	Cys	Cys	Leu	Ala	Thr	Asn	Ser	His	Leu	Gln	Ala	Glu	
		35					40					45				
CTG	ACC	TTA	GAA	GGT	CTT	GAT	AAG	ATG	ATG	CAG	CCC	GAG	CCC	ACC	TTT	192
Leu	Thr	Leu	Glu	Gly	Leu	Asp	Lys	Met	Met	Gln	Pro	Glu	Pro	Thr	Phe	
	50					55					60					
TTC	GCC	TGC	AGA	GCG	ATA	CGC	AGA	CTA	CTC	CTG	GGG	GAA	CGC	CTC	CAC	240
Phe	Ala	Cys	Arg	Ala	Ile	Arg	Arg	Leu	Leu	Leu	Gly	Glu	Arg	Leu	His	
	65				70					75					80	
CCT	TTT	ATA	CAT	CAA	GAA	GGG	ACT	CTT	TTG	GGA	AAA	GTG	GGT	CGA	CGG	288
Pro	Phe	Ile	His	Gln	Glu	Gly	Thr	Leu	Leu	Gly	Lys	Val	Gly	Arg	Arg	
				85					90					95		
TAC	AGC	GGC	GAA	GGT	TTA	ATA	ATT	GAC	GGT	GGT	GGA	GTG	TTT	ACG	CGC	336
Tyr	Ser	Gly	Glu	Gly	Leu	Ile	Ile	Asp	Gly	Gly	Gly	Val	Phe	Thr	Arg	
			100					105					110			
GGA	CAG	ATA	GAC	ACC	GAC	AAC	TAC	CTA	CCT	GCG	GTG	GGA	TCA	TGG	GAA	384
Gly	Gln	Ile	Asp	Thr	Asp	Asn	Tyr	Leu	Pro	Ala	Val	Gly	Ser	Trp	Glu	
		115					120					125				
CTT	ACC	GAT	GAT	TGT	GAT	AAA	CCC	TGC	GAA	TTC	AGG	GAG	CTA	CGC	TCG	432
Leu	Thr	Asp	Asp	Cys	Asp	Lys	Pro	Cys	Glu	Phe	Arg	Glu	Leu	Arg	Ser	
	130					135					140					
CTG	TAT	CTT	CCC	GCG	CTA	CTA	ACG	TGC	ACC	ATA	TGT	TAC	AAA	GCC	ATG	480
Leu	Tyr	Leu	Pro	Ala	Leu	Leu	Thr	Cys	Thr	Ile	Cys	Tyr	Lys	Ala	Met	
	145				150					155					160	
TTC	AGG	ATA	GTG	TGC	AGG	TAC	CTG	GAG	TTC	TGG	GAG	TTC	GAA	CAG	TGT	528
Phe	Arg	Ile	Val	Cys	Arg	Tyr	Leu	Glu	Phe	Trp	Glu	Phe	Glu	Gln	Cys	
				165				170						175		
TTT	CAT	GCG	TTT	CTG	GCG	GTG	TTG	CCC	CAT	AGT	CTA	CAA	CCC	ACA	ATC	576
Phe	His	Ala	Phe	Leu	Ala	Val	Leu	Pro	His	Ser	Leu	Gln	Pro	Thr	Ile	
			180					185					190			
TAT	CAA	AAT	TAT	TTT	GCA	CTC	CTG	GAG	AGC	CTG	AAG	CAT	CTC	TCG	TTT	624
Tyr	Gln	Asn	Tyr	Phe	Ala	Leu	Leu	Glu	Ser	Leu	Lys	His	Leu	Ser	Phe	
		195				200						205				
TCA	ATA	ATG	CCA	CCC	GCA	TCC	CCA	GAC	GCA	CAG	CTA	CAT	TTT	TTA	AAG	672
Ser	Ile	Met	Pro	Pro	Ala	Ser	Pro	Asp	Ala	Gln	Leu	His	Phe	Leu	Lys	
	210					215					220					
TTT	AAC	ATC	AGC	AGC	TTC	ATG	GCC	ACG	TGG	GGG	TGG	CAC	GGA	GAG	CTG	720
Phe	Asn	Ile	Ser	Ser	Phe	Met	Ala	Thr	Trp	Gly	Trp	His	Gly	Glu	Leu	
	225				230					235					240	
GTC	TCG	CTG	CGC	CGT	GCC	ATC	GCT	CAC	AAC	GTA	GAG	CGA	CTG	CCC	ACC	768
Val	Ser	Leu	Arg	Arg	Ala	Ile	Ala	His	Asn	Val	Glu	Arg	Leu	Pro	Thr	
				245					250					255		
GTG	CTG	AAG	AAC	CTG	TCG	AAA	CAG	AGT	AAG	CAC	CAG	GAC	GTC	AAG	GTT	816

0050749-66900

Val	Leu	Lys	Asn	Leu	Ser	Lys	Gln	Ser	Lys	His	Gln	Asp	Val	Lys	Val	
			260					265					270			
AAC	GGA	CGG	GAT	CTG	GTG	GGC	TTT	CAG	CTG	GCT	CTA	AAC	CAG	CTC	GTG	864
Asn	Gly	Arg	Asp	Leu	Val	Gly	Phe	Gln	Leu	Ala	Leu	Asn	Gln	Leu	Val	
		275					280					285				
TCC	CGT	CTG	CAC	GTA	AAA	ATC	CAA	CGC	AAG	GAC	CCC	GGA	CCA	AAG	CCA	912
Ser	Arg	Leu	His	Val	Lys	Ile	Gln	Arg	Lys	Asp	Pro	Gly	Pro	Lys	Pro	
		290				295					300					
TAC	AGG	GTG	GTC	GTC	AGT	ACC	CCA	GAT	TGT	ACC	TAC	TAT	CTA	GTG	TAT	960
Tyr	Arg	Val	Val	Val	Ser	Thr	Pro	Asp	Cys	Thr	Tyr	Tyr	Leu	Val	Tyr	
		305			310					315					320	
CCG	GGC	ACA	CCG	GCC	ATC	TAC	AGA	CTC	GTC	ATG	TGT	ATG	GCA	GTG	GCA	1008
Pro	Gly	Thr	Pro	Ala	Ile	Tyr	Arg	Leu	Val	Met	Cys	Met	Ala	Val	Ala	
				325					330					335		
GAC	TGC	ATC	GGC	CAC	TCG	TGC	AGC	GGA	CTG	CAC	CCC	TGC	GCA	AAC	TTT	1056
Asp	Cys	Ile	Gly	His	Ser	Cys	Ser	Gly	Leu	His	Pro	Cys	Ala	Asn	Phe	
			340					345					350			
TTA	GGC	ACC	CAC	GAG	ACA	CCG	CGT	CTC	CTG	GCG	GCG	ACG	CTT	TCA	AGA	1104
Leu	Gly	Thr	His	Glu	Thr	Pro	Arg	Leu	Leu	Ala	Ala	Thr	Leu	Ser	Arg	
		355					360					365				
ATC	CGG	TAC	GCG	CCG	AAA	GAC	CGG	CGA	GCA	GCC	ATG	AAA	GGA	AAT	TTG	1152
Ile	Arg	Tyr	Ala	Pro	Lys	Asp	Arg	Arg	Ala	Ala	Met	Lys	Gly	Asn	Leu	
		370				375					380					
CAG	GCG	TGC	TTC	CAA	CGA	TAC	GCG	GCC	ACG	GAC	GCG	CGG	ACT	CTG	GGC	1200
Gln	Ala	Cys	Phe	Gln	Arg	Tyr	Ala	Ala	Thr	Asp	Ala	Arg	Thr	Leu	Gly	
				390						395					400	
AGC	TCT	ACA	GTG	TCA	GAC	ATG	CTG	GAA	CCC	ACA	AAA	CAC	GTC	AGT	TTG	1248
Ser	Ser	Thr	Val	Ser	Asp	Met	Leu	Glu	Pro	Thr	Lys	His	Val	Ser	Leu	
				405					410					415		
GAA	AAC	TTC	AAG	ATC	ACC	ATA	TTC	AAC	ACC	AAC	ATG	GTG	ATT	AAC	ACT	1296
Glu	Asn	Phe	Lys	Ile	Thr	Ile	Phe	Asn	Thr	Asn	Met	Val	Ile	Asn	Thr	
			420					425					430			
AAG	ATA	AGC	TGC	CAC	GTT	CCT	AAC	ACC	CTG	CAA	AAG	ACT	ATT	TTA	AAC	1344
Lys	Ile	Ser	Cys	His	Val	Pro	Asn	Thr	Leu	Gln	Lys	Thr	Ile	Leu	Asn	
		435					440					445				
ATC	CCC	AGA	TTG	ACC	AAC	AAT	TTT	GTT	ATA	CGA	AAG	TAC	TCC	GTA	AAG	1392
Ile	Pro	Arg	Leu	Thr	Asn	Asn	Phe	Val	Ile	Arg	Lys	Tyr	Ser	Val	Lys	
		450				455					460					
GAA	CCT	TCT	TTT	ACC	ATA	AGC	GTG	TTT	TTT	TCC	GAC	AAC	ATG	TGT	CAA	1440
Glu	Pro	Ser	Phe	Thr	Ile	Ser	Val	Phe	Phe	Ser	Asp	Asn	Met	Cys	Gln	
				470						475					480	
GGC	ACC	GCA	ATA	AAC	ATC	AAC	ATC	AGT	GGG	GAC	ATG	CTG	CAC	TTT	CTC	1488
Gly	Thr	Ala	Ile	Asn	Ile	Asn	Ile	Ser	Gly	Asp	Met	Leu	His	Phe	Leu	
				485					490					495		
TTC	GCA	ATG	GGT	ACG	CTG	AAA	TGC	TTT	CTG	CCA	ATC	AGG	CAC	ATA	TTT	1536
Phe	Ala	Met	Gly	Thr	Leu	Lys	Cys	Phe	Leu	Pro	Ile	Arg	His	Ile	Phe	
			500					505					510			
CCT	GTA	TCG	ATA	GCA	AAT	TGG	AAC	TCC	ACG	TTG	GAC	CTG	CAC	GGA	CTG	1584
Pro	Val	Ser	Ile	Ala	Asn	Trp	Asn	Ser	Thr	Leu	Asp	Leu	His	Gly	Leu	
		515					520					525				

222

GAA	AAC	CAG	TAC	ATG	GTG	AGA	ATG	GGG	CGA	AAA	AAC	GTA	TTT	TGG	ACC	1632
Glu	Asn	Gln	Tyr	Met	Val	Arg	Met	Gly	Arg	Lys	Asn	Val	Phe	Trp	Thr	
530						535					540					
ACA	AAC	TTT	CCA	TCT	GTG	GTC	TCC	AGC	AAG	GAT	GGG	CTA	AAC	GTG	TCC	1680
Thr	Asn	Phe	Pro	Ser	Val	Val	Ser	Ser	Lys	Asp	Gly	Leu	Asn	Val	Ser	
545					550				555						560	
TGG	TTT	AAG	GCC	GCG	ACA	GCC	ACG	ATT	TCT	AAA	GTG	TAC	GGG	CAG	CCT	1728
Trp	Phe	Lys	Ala	Ala	Thr	Ala	Thr	Ile	Ser	Lys	Val	Tyr	Gly	Gln	Pro	
				565					570					575		
CTT	GTG	GAA	CAG	ATT	CGC	CAC	GAG	CTG	GCG	CCC	ATT	CTC	ACG	GAC	CAG	1776
Leu	Val	Glu	Gln	Ile	Arg	His	Glu	Leu	Ala	Pro	Ile	Leu	Thr	Asp	Gln	
			580					585					590			
CAC	GCG	CGC	ATC	GAC	GGA	AAC	AAA	AAT	AGA	ATA	TTC	TCC	CTA	CTT	GAG	1824
His	Ala	Arg	Ile	Asp	Gly	Asn	Lys	Asn	Arg	Ile	Phe	Ser	Leu	Leu	Glu	
		595					600					605				
CAC	AGA	AAC	CGT	TCC	CAA	ATA	CAG	ACG	CTA	CAC	AAA	AGG	TTC	CTG	GAG	1872
His	Arg	Asn	Arg	Ser	Gln	Ile	Gln	Thr	Leu	His	Lys	Arg	Phe	Leu	Glu	
	610					615					620					
TGT	CTG	GTG	GAA	TGC	TGT	TCG	TTT	CTC	AGG	CTT	GAC	GTG	GCT	TGC	ATT	1920
Cys	Leu	Val	Glu	Cys	Cys	Ser	Phe	Leu	Arg	Leu	Asp	Val	Ala	Cys	Ile	
625				630						635					640	
AGG	CGA	GCC	GCC	GCC	CGG	GGC	CTG	TTT	GAC	TTC	TCA	AAG	AAG	ATA	ATC	1968
Arg	Arg	Ala	Ala	Ala	Arg	Gly	Leu	Phe	Asp	Phe	Ser	Lys	Lys	Ile	Ile	
				645					650					655		
AGT	CAC	ACT	AAA	AGC	AAA	CAC	GAG	TGC	GCA	GTA	CTG	GGA	TAT	AAA	AAG	2016
Ser	His	Thr	Lys	Ser	Lys	His	Glu	Cys	Ala	Val	Leu	Gly	Tyr	Lys	Lys	
			660					665					670			
TGT	AAC	CTA	ATC	CCG	AAA	ATC	TAT	GCC	CGA	AAC	AAG	AAG	ACC	AGG	CTA	2064
Cys	Asn	Leu	Ile	Pro	Lys	Ile	Tyr	Ala	Arg	Asn	Lys	Lys	Thr	Arg	Leu	
		675					680					685				
GAC	GAG	TTG	GGC	CGC	AAT	GCA	AAC	TTC	ATT	TCG	TTC	GTC	GCC	ACC	ACG	2112
Asp	Glu	Leu	Gly	Arg	Asn	Ala	Asn	Phe	Ile	Ser	Phe	Val	Ala	Thr	Thr	
	690					695					700					
GGT	CAT	CGG	TTC	GCC	GCT	CTA	AAG	CCA	CAA	ATT	GTC	CGT	CAC	GCC	ATT	2160
Gly	His	Arg	Phe	Ala	Ala	Leu	Lys	Pro	Gln	Ile	Val	Arg	His	Ala	Ile	
705					710					715					720	
CGC	AAA	CTA	GGC	CTG	CAC	TGG	CGC	CAC	CGA	ACG	GCC	GCG	TCC	AAC	GAG	2208
Arg	Lys	Leu	Gly	Leu	His	Trp	Arg	His	Arg	Thr	Ala	Ala	Ser	Asn	Glu	
				725					730					735		
CAG	ACA	CCG	CCA	GCC	GAT	CCC	CGC	GTA	CGT	TGC	GTC	CGT	CCG	CTG	GTC	2256
Gln	Thr	Pro	Pro	Ala	Asp	Pro	Arg	Val	Arg	Cys	Val	Arg	Pro	Leu	Val	
			740					745					750			
TAA																2259

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 752 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

006290 6240960

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met 1	Ala	Ala	Leu	Glu 5	Gly	Pro	Leu	Leu	Leu 10	Pro	Pro	Ser	Ala	Ser 15	Leu
Thr	Thr	Ser	Pro 20	Gln	Thr	Thr	Cys	Tyr 25	Gln	Ala	Thr	Trp	Glu 30	Ser	Gln
Leu	Glu	Ile 35	Phe	Cys	Cys	Leu	Ala 40	Thr	Asn	Ser	His	Leu 45	Gln	Ala	Glu
Leu	Thr 50	Leu	Glu	Gly	Leu	Asp 55	Lys	Met	Met	Gln	Pro 60	Glu	Pro	Thr	Phe
Phe 65	Ala	Cys	Arg	Ala	Ile 70	Arg	Arg	Leu	Leu	Leu 75	Gly	Glu	Arg	Leu	His 80
Pro	Phe	Ile	His	Gln 85	Glu	Gly	Thr	Leu	Leu 90	Gly	Lys	Val	Gly	Arg 95	Arg
Tyr	Ser	Gly	Glu 100	Gly	Leu	Ile	Ile	Asp 105	Gly	Gly	Gly	Val	Phe 110	Thr	Arg
Gly	Gln	Ile 115	Asp	Thr	Asp	Asn	Tyr 120	Leu	Pro	Ala	Val	Gly 125	Ser	Trp	Glu
Leu	Thr 130	Asp	Asp	Cys	Asp	Lys 135	Pro	Cys	Glu	Phe	Arg 140	Glu	Leu	Arg	Ser
Leu 145	Tyr	Leu	Pro	Ala	Leu 150	Leu	Thr	Cys	Thr	Ile 155	Cys	Tyr	Lys	Ala	Met 160
Phe	Arg	Ile	Val	Cys 165	Arg	Tyr	Leu	Glu	Phe 170	Trp	Glu	Phe	Glu	Gln 175	Cys
Phe	His	Ala	Phe 180	Leu	Ala	Val	Leu	Pro 185	His	Ser	Leu	Gln	Pro 190	Thr	Ile
Tyr	Gln	Asn 195	Tyr	Phe	Ala	Leu 200	Leu	Glu	Ser	Leu	Lys	His 205	Leu	Ser	Phe
Ser	Ile 210	Met	Pro	Pro	Ala	Ser 215	Pro	Asp	Ala	Gln	Leu 220	His	Phe	Leu	Lys
Phe 225	Asn	Ile	Ser	Ser	Phe 230	Met	Ala	Thr	Trp	Gly 235	Trp	His	Gly	Glu	Leu 240
Val	Ser	Leu	Arg	Arg 245	Ala	Ile	Ala	His	Asn 250	Val	Glu	Arg	Leu	Pro 255	Thr
Val	Leu	Lys	Asn 260	Leu	Ser	Lys	Gln	Ser 265	Lys	His	Gln	Asp	Val 270	Lys	Val
Asn	Gly	Arg 275	Asp	Leu	Val	Gly	Phe 280	Gln	Leu	Ala	Leu	Asn 285	Gln	Leu	Val
Ser	Arg 290	Leu	His	Val	Lys	Ile 295	Gln	Arg	Lys	Asp	Pro 300	Gly	Pro	Lys	Pro
Tyr 305	Arg	Val	Val	Val	Ser 310	Thr	Pro	Asp	Cys	Thr 315	Tyr	Tyr	Leu	Val	Tyr 320
Pro	Gly	Thr	Pro	Ala 325	Ile	Tyr	Arg	Leu	Val 330	Met	Cys	Met	Ala	Val 335	Ala

225

690		695		700
Gly 705	His Arg Phe Ala Ala	Leu Lys Pro Gln 710	Ile Val Arg His Ala	Ile 720
Arg	Lys Leu Gly Leu 725	His Trp Arg His 730	Arg Thr Ala Ala Ser Asn 735	Glu
Gln Thr Pro 740	Pro Ala Asp Pro Arg Val 745	Arg Cys Val Arg Pro Leu Val 750		

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 364 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..364
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATG GTA CGT CCA ACC GAG GCC GAG GTT AAG AAA TCC CTG AGC AGG CTT	48
Met Val Arg Pro Thr Glu Ala Glu Val Lys Lys Ser Leu Ser Arg Leu	
1 5 10 15	
CCA GCA GCA CGC AAA AGA GCA GGT AAC CGG GCC CAC CTG GCC ACC TAC	96
Pro Ala Ala Arg Lys Arg Ala Gly Asn Arg Ala His Leu Ala Thr Tyr	
20 25 30	
CGC CGG CTC CTC AAG TAC TCC ACC CTG CCC GAT CTA TGG CGG TTT CTA	144
Arg Arg Leu Leu Lys Tyr Ser Thr Leu Pro Asp Leu Trp Arg Phe Leu	
35 40 45	
AGT AGC CGG CCC CAG AAC CCT CCC CTT GGA CAC CAC AGA TTA TTC TTT	192
Ser Ser Arg Pro Gln Asn Pro Pro Leu Gly His His Arg Leu Phe Phe	
50 55 60	
GAG GTG ACT CTA GGG CAC AGA ATT GCC GAC TGC GTA ATT CTG GTA TCG	240
Glu Val Thr Leu Gly His Arg Ile Ala Asp Cys Val Ile Leu Val Ser	
65 70 75 80	
GGT GGG CAT CAG CCC GTA TGT TAC GTT GTA GAG CTC AAG ACT TGT CTG	288
Gly Gly His Gln Pro Val Cys Tyr Val Val Glu Leu Lys Thr Cys Leu	
85 90 95	
AGT CAC CAG CTG ATC CCA ACC AAC ACC GTG AGA ACG TCA CAG CGA GCT	336
Ser His Gln Leu Ile Pro Thr Asn Thr Val Arg Thr Ser Gln Arg Ala	
100 105 110	
CAA GGC CTG TGC CAA CTC TCC GAC TCG A	364
Gln Gly Leu Cys Gln Leu Ser Asp Ser	
115 120	

006290-624990

226

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Val Arg Pro Thr Glu Ala Glu Val Lys Lys Ser Leu Ser Arg Leu
 1 5 10 15
 Pro Ala Ala Arg Lys Arg Ala Gly Asn Arg Ala His Leu Ala Thr Tyr
 20 25 30
 Arg Arg Leu Leu Lys Tyr Ser Thr Leu Pro Asp Leu Trp Arg Phe Leu
 35 40 45
 Ser Ser Arg Pro Gln Asn Pro Pro Leu Gly His His Arg Leu Phe Phe
 50 55 60
 Glu Val Thr Leu Gly His Arg Ile Ala Asp Cys Val Ile Leu Val Ser
 65 70 75 80
 Gly Gly His Gln Pro Val Cys Tyr Val Val Glu Leu Lys Thr Cys Leu
 85 90 95
 Ser His Gln Leu Ile Pro Thr Asn Thr Val Arg Thr Ser Gln Arg Ala
 100 105 110
 Gln Gly Leu Cys Gln Leu Ser Asp Ser
 115 120

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..918
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATG GCA CTC GAC AAG AGT ATA GTG GTT AAC TTC ACC TCC AGA CTC TTC
 Met Ala Leu Asp Lys Ser Ile Val Val Asn Phe Thr Ser Arg Leu Phe
 1 5 10 15 48
 GCT GAT GAA CTG GCC GCC CTT CAG TCA AAA ATA GGG AGC GTA CTG CCG
 Ala Asp Glu Leu Ala Ala Leu Gln Ser Lys Ile Gly Ser Val Leu Pro
 20 25 30 96

227

CTC	GGA	GAT	TGC	CAC	CGT	TTA	CAA	AAT	ATA	CAG	GCA	TTG	GGC	CTG	GGG	144
Leu	Gly	Asp	Cys	His	Arg	Leu	Gln	Asn	Ile	Gln	Ala	Leu	Gly	Leu	Gly	
		35					40					45				
TGC	GTA	TGC	TCA	CGT	GAG	ACA	TCT	CCG	GAC	TAC	ATC	CAA	ATT	ATG	CAG	192
Cys	Val	Cys	Ser	Arg	Glu	Thr	Ser	Pro	Asp	Tyr	Ile	Gln	Ile	Met	Gln	
	50					55					60					
TAT	CTA	TCC	AAG	TGC	ACA	CTC	GCT	GTC	CTG	GAG	GAG	GTT	CGC	CCG	GAC	240
Tyr	Leu	Ser	Lys	Cys	Thr	Leu	Ala	Val	Leu	Glu	Glu	Val	Arg	Pro	Asp	
	65				70					75					80	
AGC	CTG	CGC	CTA	ACG	CGG	ATG	GAT	CCC	TCT	GAC	AAC	CTT	CAG	ATA	AAA	288
Ser	Leu	Arg	Leu		Arg	Met	Asp	Pro	Ser	Asp	Asn	Leu	Gln	Ile	Lys	
				85					90					95		
AAC	GTA	TAT	GCC	CCC	TTT	TTT	CAG	TGG	GAC	AGC	AAC	ACC	CAG	CTA	GCA	336
Asn	Val	Tyr	Ala	Pro	Phe	Phe	Gln	Trp	Asp	Ser	Asn	Thr	Gln	Leu	Ala	
			100					105					110			
GTG	CTA	CCC	CCA	TTT	TTT	AGC	CGA	AAG	GAT	TCC	ACC	ATT	GTG	CTC	GAA	384
Val	Leu	Pro	Pro	Phe	Phe	Ser	Arg	Lys	Asp	Ser	Thr	Ile	Val	Leu	Glu	
		115					120					125				
TCC	AAC	GGA	TTT	GAC	CCC	GTG	TTC	CCC	ATG	GTC	GTG	CCG	CAG	CAA	CTG	432
Ser	Asn	Gly	Phe	Asp	Pro	Val	Phe	Pro	Met	Val	Val	Pro	Gln	Gln	Leu	
	130					135					140					
GGG	CAC	GCT	ATT	CTG	CAG	CAG	CTG	TTG	GTG	TAC	CAC	ATC	TAC	TCC	AAA	480
Gly	His	Ala	Ile	Leu	Gln	Gln	Leu	Leu	Val	Tyr	His	Ile	Tyr	Ser	Lys	
	145				150					155					160	
ATA	TCG	GCC	GGG	GCC	CCG	GAT	GAT	GTA	AAT	ATG	GCG	GAA	CTT	GAT	CTA	528
Ile	Ser	Ala	Gly	Ala	Pro	Asp	Asp	Val	Asn	Met	Ala	Glu	Leu	Asp	Leu	
				165					170					175		
TAT	ACC	ACC	AAT	GTG	TCA	TTT	ATG	GGG	CGC	ACA	TAT	CGT	CTG	GAC	GTA	576
Tyr	Thr	Thr	Asn	Val	Ser	Phe	Met	Gly	Arg	Thr	Tyr	Arg	Leu	Asp	Val	
			180					185					190			
GAC	AAC	ACG	GAT	CCA	CGT	ACT	GCC	CTG	CGA	GTG	CTT	GAC	GAT	CTG	TCC	624
Asp	Asn	Thr	Asp	Pro	Arg	Thr	Ala	Leu	Arg	Val	Leu	Asp	Asp	Leu	Ser	
			195				200					205				
ATG	TAC	CTT	TGT	ATC	CTA	TCA	GCC	TTG	GTT	CCC	AGG	GGG	TGT	CTC	CGT	672
Met	Tyr	Leu	Cys	Ile	Leu	Ser	Ala	Leu	Val	Pro	Arg	Gly	Cys	Leu	Arg	
	210					215					220					
CTG	CTC	ACG	GCG	CTC	GTG	CGG	CAC	GAC	AGG	CAT	CCT	CTG	ACA	GAG	GTG	720
Leu	Leu	Thr	Ala	Leu	Val	Arg	His	Asp	Arg	His	Pro	Leu	Thr	Glu	Val	
	225				230					235					240	
TTT	GAG	GGG	GTG	GTG	CCA	GAT	GAG	GTG	ACC	AGG	ATA	GAT	CTC	GAC	CAG	768
Phe	Glu	Gly	Val	Val	Pro	Asp	Glu	Val	Thr	Arg	Ile	Asp	Leu	Asp	Gln	
				245					250					255		
TTG	AGC	GTC	CCA	GAT	GAC	ATC	ACC	AGG	ATG	CGC	GTC	ATG	TTC	TCC	TAT	816
Leu	Ser	Val	Pro	Asp	Asp	Ile	Thr	Arg	Met	Arg	Val	Met	Phe	Ser	Tyr	
			260					265					270			
CTT	CAG	AGT	CTC	AGT	TCT	ATA	TTT	AAT	CTT	GGC	CCC	AGA	CTG	CAC	GTG	864
Leu	Gln	Ser	Leu	Ser	Ser	Ile	Phe	Asn	Leu	Gly	Pro	Arg	Leu	His	Val	
			275				280					285				
TAT	GCC	TAC	TCG	GCA	GAG	ACT	TTG	GCG	GCC	TCC	TGT	TGG	TAT	TCC	CCA	912
Tyr	Ala	Tyr	Ser	Ala	Glu	Thr	Leu	Ala	Ala	Ser	Cys	Trp	Tyr	Ser	Pro	
	290					295					300					

006290-6472960

228

CGC TAA
Arg
305

918

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 305 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Ala Leu Asp Lys Ser Ile Val Val Asn Phe Thr Ser Arg Leu Phe
1 5 10 15
Ala Asp Glu Leu Ala Ala Leu Gln Ser Lys Ile Gly Ser Val Leu Pro
20 25 30
Leu Gly Asp Cys His Arg Leu Gln Asn Ile Gln Ala Leu Gly Leu Gly
35 40 45
Cys Val Cys Ser Arg Glu Thr Ser Pro Asp Tyr Ile Gln Ile Met Gln
50 55 60
Tyr Leu Ser Lys Cys Thr Leu Ala Val Leu Glu Glu Val Arg Pro Asp
65 70 75 80
Ser Leu Arg Leu Thr Arg Met Asp Pro Ser Asp Asn Leu Gln Ile Lys
85 90 95
Asn Val Tyr Ala Pro Phe Phe Gln Trp Asp Ser Asn Thr Gln Leu Ala
100 105 110
Val Leu Pro Pro Phe Phe Ser Arg Lys Asp Ser Thr Ile Val Leu Glu
115 120 125
Ser Asn Gly Phe Asp Pro Val Phe Pro Met Val Val Pro Gln Gln Leu
130 135 140
Gly His Ala Ile Leu Gln Gln Leu Leu Val Tyr His Ile Tyr Ser Lys
145 150 155 160
Ile Ser Ala Gly Ala Pro Asp Asp Val Asn Met Ala Glu Leu Asp Leu
165 170 175
Tyr Thr Thr Asn Val Ser Phe Met Gly Arg Thr Tyr Arg Leu Asp Val
180 185 190
Asp Asn Thr Asp Pro Arg Thr Ala Leu Arg Val Leu Asp Asp Leu Ser
195 200 205
Met Tyr Leu Cys Ile Leu Ser Ala Leu Val Pro Arg Gly Cys Leu Arg
210 215 220
Leu Leu Thr Ala Leu Val Arg His Asp Arg His Pro Leu Thr Glu Val
225 230 235 240
Phe Glu Gly Val Val Pro Asp Glu Val Thr Arg Ile Asp Leu Asp Gln
245 250 255
Leu Ser Val Pro Asp Asp Ile Thr Arg Met Arg Val Met Phe Ser Tyr
260 265 270

006290-0240988

229

Leu Gln Ser Leu Ser Ser Ile Phe Asn Leu Gly Pro Arg Leu His Val
 275 280 285

Tyr Ala Tyr Ser Ala Glu Thr Leu Ala Ala Ser Cys Trp Tyr Ser Pro
 290 295 300

Arg
 305

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 873 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..873
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATG GCG TCA TCT GAT ATT CTG TCG GTT GCA AGG ACG GAT GAC GGC TCC	48
Met Ala Ser Ser Asp Ile Leu Ser Val Ala Arg Thr Asp Asp Gly Ser	
1 5 10 15	
GTC TGT GAA GTC TCC CTG CGT GGA GGT AGG AAA AAA ACT ACC GTC TAC	96
Val Cys Glu Val Ser Leu Arg Gly Gly Arg Lys Lys Thr Thr Val Tyr	
20 25 30	
CTG CCG GAC ACT GAA CCC TGG GTG GTA GAG ACC GAC GCC ATC AAA GAC	144
Leu Pro Asp Thr Glu Pro Trp Val Val Glu Thr Asp Ala Ile Lys Asp	
35 40 45	
GCC TTC CTC AGC GAC GGG ATC GTG GAT ATG GCT CGA AAG CTT CAT CGT	192
Ala Phe Leu Ser Asp Gly Ile Val Asp Met Ala Arg Lys Leu His Arg	
50 55 60	
GGT GCC CTG CCC TCA AAT TCT CAC AAC GGC TTG AGG ATG GTG CTT TTT	240
Gly Ala Leu Pro Ser Asn Ser His Asn Gly Leu Arg Met Val Leu Phe	
65 70 75 80	
TGT TAT TGT TAC TTG CAA AAT TGT GTG TAC CTA GCC CTG TTT CTG TGC	288
Cys Tyr Cys Tyr Leu Gln Asn Cys Val Tyr Leu Ala Leu Phe Leu Cys	
85 90 95	
CCC CTT AAT CCT TAC TTG GTA ACT CCC TCA AGC ATT GAG TTT GCC GAG	336
Pro Leu Asn Pro Tyr Leu Val Thr Pro Ser Ser Ile Glu Phe Ala Glu	
100 105 110	
CCC GTT GTG GCA CCT GAG GTG CTC TTC CCA CAC CCG GCT GAG ATG TCT	384
Pro Val Val Ala Pro Glu Val Leu Phe Pro His Pro Ala Glu Met Ser	
115 120 125	
CGC GGT TGC GAT GAC GCG ATT TTC TGT AAA CTG CCC TAT ACC GTG CCT	432
Arg Gly Cys Asp Asp Ala Ile Phe Cys Lys Leu Pro Tyr Thr Val Pro	
130 135 140	

006290-64720960

230

ATA	ATC	AAC	ACC	ACG	TTT	GGA	CGC	ATT	TAC	CCG	AAC	TCT	ACA	CGC	GAG	480
Ile	Ile	Asn	Thr	Thr	Phe	Gly	Arg	Ile	Tyr	Pro	Asn	Ser	Thr	Arg	Glu	
145					150					155					160	
CCG	GAC	GGC	AGG	CCT	ACG	GAT	TAC	TCC	ATG	GCC	CTT	AGA	AGG	GCT	TTT	528
Pro	Asp	Gly	Arg	Pro	Thr	Asp	Tyr	Ser	Met	Ala	Leu	Arg	Arg	Ala	Phe	
				165					170					175		
GCA	GTT	ATG	GTT	AAC	ACG	TCA	TGT	GCA	GGA	GTG	ACA	TTG	TGC	CGC	GGA	576
Ala	Val	Met	Val	Asn	Thr	Ser	Cys	Ala	Gly	Val	Thr	Leu	Cys	Arg	Gly	
			180					185					190			
GAA	ACT	CAG	ACC	GCA	TCC	CGT	AAC	CAC	ACT	GAG	TGG	GAA	AAT	CTG	CTG	624
Glu	Thr	Gln	Thr	Ala	Ser	Arg	Asn	His	Thr	Glu	Trp	Glu	Asn	Leu	Leu	
		195					200					205				
GCT	ATG	TTT	TCT	GTG	ATT	ATC	TAT	GCC	TTA	GAT	CAC	AAC	TGT	CAC	CCG	672
Ala	Met	Phe	Ser	Val	Ile	Ile	Tyr	Ala	Leu	Asp	His	Asn	Cys	His	Pro	
	210					215				220						
GAA	GCA	CTG	TCT	ATC	GCG	AGC	GGC	ATC	TTT	GAC	GAG	CGT	GAC	TAT	GGA	720
Glu	Ala	Leu	Ser	Ile	Ala	Ser	Gly	Ile	Phe	Asp	Glu	Arg	Asp	Tyr	Gly	
225					230					235					240	
TTA	TTC	ATC	TCT	CAG	CCC	CGG	AGC	GTG	CCC	TCG	CCT	ACC	CCT	TGC	GAC	768
Leu	Phe	Ile	Ser	Gln	Pro	Arg	Ser	Val	Pro	Ser	Pro	Thr	Pro	Cys	Asp	
				245					250					255		
GTG	TCG	TGG	GAA	GAT	ATC	TAC	AAC	GGG	ACT	TAC	CTA	GCT	CGG	CCT	GGA	816
Val	Ser	Trp	Glu	Asp	Ile	Tyr	Asn	Gly	Thr	Tyr	Leu	Ala	Arg	Pro	Gly	
			260					265					270			
AAC	TGT	GAC	CCC	TGG	CCC	AAT	CTA	TCC	ACC	CCT	CCC	TTG	ATT	CTA	AAT	864
Asn	Cys	Asp	Pro	Trp	Pro	Asn	Leu	Ser	Thr	Pro	Pro	Leu	Ile	Leu	Asn	
		275					280					285				
TTT	AAA	TAA														873
Phe	Lys															
	290															

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met	Ala	Ser	Ser	Asp	Ile	Leu	Ser	Val	Ala	Arg	Thr	Asp	Asp	Gly	Ser
1				5					10					15	
Val	Cys	Glu	Val	Ser	Leu	Arg	Gly	Gly	Arg	Lys	Lys	Thr	Thr	Val	Tyr
			20					25					30		
Leu	Pro	Asp	Thr	Glu	Pro	Trp	Val	Val	Glu	Thr	Asp	Ala	Ile	Lys	Asp
			35				40					45			
Ala	Phe	Leu	Ser	Asp	Gly	Ile	Val	Asp	Met	Ala	Arg	Lys	Leu	His	Arg
			50			55					60				
Gly	Ala	Leu	Pro	Ser	Asn	Ser	His	Asn	Gly	Leu	Arg	Met	Val	Leu	Phe
65					70				75						80

231

Cys Tyr Cys Tyr Leu Gln Asn Cys Val Tyr Leu Ala Leu Phe Leu Cys
 85 90 95
 Pro Leu Asn Pro Tyr Leu Val Thr Pro Ser Ser Ile Glu Phe Ala Glu
 100 105 110
 Pro Val Val Ala Pro Glu Val Leu Phe Pro His Pro Ala Glu Met Ser
 115 120 125
 Arg Gly Cys Asp Asp Ala Ile Phe Cys Lys Leu Pro Tyr Thr Val Pro
 130 135 140
 Ile Ile Asn Thr Thr Phe Gly Arg Ile Tyr Pro Asn Ser Thr Arg Glu
 145 150 155 160
 Pro Asp Gly Arg Pro Thr Asp Tyr Ser Met Ala Leu Arg Arg Ala Phe
 165 170 175
 Ala Val Met Val Asn Thr Ser Cys Ala Gly Val Thr Leu Cys Arg Gly
 180 185 190
 Glu Thr Gln Thr Ala Ser Arg Asn His Thr Glu Trp Glu Asn Leu Leu
 195 200 205
 Ala Met Phe Ser Val Ile Ile Tyr Ala Leu Asp His Asn Cys His Pro
 210 215 220
 Glu Ala Leu Ser Ile Ala Ser Gly Ile Phe Asp Glu Arg Asp Tyr Gly
 225 230 235 240
 Leu Phe Ile Ser Gln Pro Arg Ser Val Pro Ser Pro Thr Pro Cys Asp
 245 250 255
 Val Ser Trp Glu Asp Ile Tyr Asn Gly Thr Tyr Leu Ala Arg Pro Gly
 260 265 270
 Asn Cys Asp Pro Trp Pro Asn Leu Ser Thr Pro Pro Leu Ile Leu Asn
 275 280 285
 Phe Lys
 290

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 363 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..363
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATG AGC ATG ACT TTC CCC GTC TCC AGT CAC CGG AGG AAT GGT GGA CGG
 Met Ser Met Thr Phe Pro Val Ser Ser His Arg Arg Asn Gly Gly Arg
 1 5 10 15

232

CTC CGT CCT GGT GCG AAT GGC CAC CAA GCC TCC CGT GAT TGG TCT TAT 96
 Leu Arg Pro Gly Ala Asn Gly His Gln Ala Ser Arg Asp Trp Ser Tyr
 20 25 30

AAC AGT GCT CTT CCT CCT AGT CAT AGG CGC CTG CGT CTA CTG CTG CAT 144
 Asn Ser Ala Leu Pro Pro Ser His Arg Arg Leu Arg Leu Leu His
 35 40 45

TCG CGT GTT CCT GGC GGC TCG ACT GTG GCG CGC CAC CCC ACT AGG CAG 192
 Ser Arg Val Pro Gly Gly Ser Thr Val Ala Arg His Pro Thr Arg Gln
 50 55 60

GGC CAC CGT GGC GTA TCA GGT CCT TCG CAC CCT GGG ACC GCA GGC CGG 240
 Gly His Arg Gly Val Ser Gly Pro Ser His Pro Gly Thr Ala Gly Arg
 65 70 75 80

GTC ACA TGC ACC GCC GAC GGT GGG CAT AGC TAC CCA GGA GCC CTA CCG 288
 Val Thr Cys Thr Ala Asp Gly Gly His Ser Tyr Pro Gly Ala Leu Pro
 85 90 95

TAC AAT ATA CAT GCC AGA TTA GAA CGG GGT GTG TGC TAT AAT GGA TGG 336
 Tyr Asn Ile His Ala Arg Leu Glu Arg Gly Val Cys Tyr Asn Gly Trp
 100 105 110

CTA TGG GGG GGG GCT GTA GAT AAT TGA 363
 Leu Trp Gly Gly Ala Val Asp Asn
 115 120

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Ser Met Thr Phe Pro Val Ser Ser His Arg Arg Asn Gly Gly Arg
 5 10 15

Leu Arg Pro Gly Ala Asn Gly His Gln Ala Ser Arg Asp Trp Ser Tyr
 20 25 30

Asn Ser Ala Leu Pro Pro Ser His Arg Arg Leu Arg Leu Leu His
 35 40 45

Ser Arg Val Pro Gly Gly Ser Thr Val Ala Arg His Pro Thr Arg Gln
 50 55 60

Gly His Arg Gly Val Ser Gly Pro Ser His Pro Gly Thr Ala Gly Arg
 65 70 75 80

Val Thr Cys Thr Ala Asp Gly Gly His Ser Tyr Pro Gly Ala Leu Pro
 85 90 95

Tyr Asn Ile His Ala Arg Leu Glu Arg Gly Val Cys Tyr Asn Gly Trp
 100 105 110

Leu Trp Gly Gly Ala Val Asp Asn
 115 120

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: N

(A) NAME/KEY: CDS

(B) LOCATION: 1..921

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATG Met 1	CTG Leu	CTC Leu	AGC Ser	ACG Arg 5	CAC His	AGG Arg	GAG Glu	CGC Arg	CTT Leu 10	GCC Ala	GCC Ala	AAC Asn	CTG Leu	GAG Glu 15	GAG Glu	48
ACC Thr	GCC Ala	AAA Lys	GAC Asp 20	GCC Ala	GGA Gly	GAG Glu	AGG Arg	TGG Trp 25	GAA Glu	CTG Leu	AGT Ser	GCC Ala	CCG Pro 30	ACA Thr	TTC Phe	96
ACG Thr	CGA Arg	CAC His 35	TGT Cys	CCC Pro	AAA Lys	ACG Thr	GCA Ala 40	CGG Arg	ATG Met	GCG Ala	CAC His	CCT Pro 45	TTT Phe	ATT Ile	GGC Gly	144
GTG Val 50	GTG Val	CAC His	AGA Arg	ATA Ile	AAC Asn	TCA Ser 55	TAC Tyr	AGT Ser	TCG Ser	GTC Val	CTG Leu 60	GAA Glu	ACA Thr	TAC Tyr	TGC Cys	192
ACA Thr 65	CGG Arg	CAC His	CAT His	CCC Pro	GCC Ala 70	ACG Thr	CCC Pro	ACG Thr	TCA Ser	GCA Ala 75	AAT Asn	CCC Pro	GAC Asp	GTG Val	GGA Gly 80	240
ACC Thr	CCC Pro	AGA Arg	CCG Pro	TCC Ser 85	GAG Glu	GAC Asp	AAC Asn	GTC Val	CCC Pro 90	GCA Ala	AAG Lys	CCG Pro	CGC Arg	CTA Leu 95	TTG Leu	288
GAG Glu	TCC Ser	CTA Leu	TCA Ser 100	ACA Thr	TAC Tyr	TTG Leu	CAG Gln	ATG Met 105	CGG Arg	TGT Cys	GTG Val	CGC Arg	GAG Glu 110	GAC Asp	GCG Ala	336
CAC His	GTC Val	TCC Ser 115	ACG Thr	GCC Ala	GAT Asp	CAA Gln	CTG Leu 120	GTC Val	GAG Glu	TAC Tyr	CAG Gln	GCG Ala 125	GGC Gly	AGA Arg	AAA Lys	384
ACA Thr 130	CAC His	GAC Asp	TCC Ser	CTG Leu	CAC His	GCC Ala 135	TGC Cys	TCT Ser	GTC Val	TAC Tyr	CGC Arg	GAA Glu 140	CTT Leu	CAG Gln	GCT Ala	432
TTT Phe 145	CTG Leu	GTT Val	AAC Asn	CTT Leu 150	TCG Ser	TCC Ser	TTT Phe	CTG Leu	AAC Asn	GGC Gly 155	TGT Cys	TAC Tyr	GTT Val	CCC Pro	GGG Gly 160	480
GTG Val	CAC His	TGG Trp	CTG Leu	GAG Glu 165	CCC Pro	TTC Phe	CAA Gln	CAG Gln	CAG Gln 170	CTA Leu	GTA Val	ATG Met	CAC His	ACT Thr 175	TTT Phe	528
TTC Phe	TTT Phe	TTG Leu	GTT Val 180	TCA Ser	ATC Ile	AAG Lys	GCC Ala	CCA Pro 185	CAA Gln	AAG Lys	ACG Thr	CAC His	CAG Gln 190	TTG Leu	TTT Phe	576

234

GGA TTT TTT AAG CAG TAC TTC GGT TTA TTT GAA ACT CCA AAC AGT GTT	624
Gly Leu Phe Lys Gln Tyr Phe Gly Leu Phe Glu Thr Pro Asn Ser Val	
195 200 205	
TTA CAG ACG TTT AAG CAA AAG GCA AGC GTA TTC CTA ATA CCA AGG AGA	672
Leu Gln Thr Phe Lys Gln Lys Ala Ser Val Phe Leu Ile Pro Arg Arg	
210 215 220	
CAC GGA AAG ACA TGG ATA GTG GTG GCG ATC ATC AGC ATG CTA CTG GCA	720
His Gly Lys Thr Trp Ile Val Val Ala Ile Ile Ser Met Leu Leu Ala	
225 230 235 240	
TCC GTA GAG AAC ATT AAC ATT GGG TAC GTA GCC CAC CAA AAG CAC GTA	768
Ser Val Glu Asn Ile Asn Ile Gly Tyr Val Ala His Gln Lys His Val	
245 250 255	
GCC AAC TCC GTG TTC GCG GAA ATC ATA AAG ACG CTT TGT CGG TGG TTC	816
Ala Asn Ser Val Phe Ala Glu Ile Ile Lys Thr Leu Cys Arg Trp Phe	
260 265 270	
CCC CCC AAA AAT TTA AAC ATC AAG AAG GAG AAC GGA ACC ATA ATC TAC	864
Pro Pro Lys Asn Leu Asn Ile Lys Lys Glu Asn Gly Thr Ile Ile Tyr	
275 280 285	
ACG CGA CCC GGA GGA CGG TCC AGC TCG CTG ATG TGC GCA ACA TGC TTC	912
Thr Arg Pro Gly Gly Arg Ser Ser Ser Leu Met Cys Ala Thr Cys Phe	
290 295 300	
AAT AAG AAC	921
Asn Lys Asn	
305	

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 307 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met	Leu	Leu	Ser	Arg	His	Arg	Glu	Arg	Leu	Ala	Ala	Asn	Leu	Glu	Glu
1				5					10					15	
Thr	Ala	Lys	Asp	Ala	Gly	Glu	Arg	Trp	Glu	Leu	Ser	Ala	Pro	Thr	Phe
		20						25					30		
Thr	Arg	His	Cys	Pro	Lys	Thr	Ala	Arg	Met	Ala	His	Pro	Phe	Ile	Gly
		35					40					45			
Val	Val	His	Arg	Ile	Asn	Ser	Tyr	Ser	Ser	Val	Leu	Glu	Thr	Tyr	Cys
	50					55					60				
Thr	Arg	His	His	Pro	Ala	Thr	Pro	Thr	Ser	Ala	Asn	Pro	Asp	Val	Gly
65				70					75					80	
Thr	Pro	Arg	Pro	Ser	Glu	Asp	Asn	Val	Pro	Ala	Lys	Pro	Arg	Leu	Leu
			85					90						95	
Glu	Ser	Leu	Ser	Thr	Tyr	Leu	Gln	Met	Arg	Cys	Val	Arg	Glu	Asp	Ala
		100						105					110		
His	Val	Ser	Thr	Ala	Asp	Gln	Leu	Val	Glu	Tyr	Gln	Ala	Gly	Arg	Lys
		115					120						125		

235

Thr His Asp Ser Leu His Ala Cys Ser Val Tyr Arg Glu Leu Gln Ala
 130 135 140
 Phe Leu Val Asn Leu Ser Ser Phe Leu Asn Gly Cys Tyr Val Pro Gly
 145 150 155 160
 Val His Trp Leu Glu Pro Phe Gln Gln Gln Leu Val Met His Thr Phe
 165 170 175
 Phe Phe Leu Val Ser Ile Lys Ala Pro Gln Lys Thr His Gln Leu Phe
 180 185 190
 Gly Leu Phe Lys Gln Tyr Phe Gly Leu Phe Glu Thr Pro Asn Ser Val
 195 200 205
 Leu Gln Thr Phe Lys Gln Lys Ala Ser Val Phe Leu Ile Pro Arg Arg
 210 215 220
 His Gly Lys Thr Trp Ile Val Val Ala Ile Ile Ser Met Leu Leu Ala
 225 230 235 240
 Ser Val Glu Asn Ile Asn Ile Gly Tyr Val Ala His Gln Lys His Val
 245 250 255
 Ala Asn Ser Val Phe Ala Glu Ile Ile Lys Thr Leu Cys Arg Trp Phe
 260 265 270
 Pro Pro Lys Asn Leu Asn Ile Lys Lys Glu Asn Gly Thr Ile Ile Tyr
 275 280 285
 Thr Arg Pro Gly Gly Arg Ser Ser Ser Leu Met Cys Ala Thr Cys Phe
 290 295 300
 Asn Lys Asn
 305

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1365
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG GAT GCG CAT GCT ATC AAC GAA AGA TAC GTA GGT CCT CGC TGC CAC	48
Met Asp Ala His Ala Ile Asn Glu Arg Tyr Val Gly Pro Arg Cys His	
1 5 10 15	
CGT TTG GCC CAC GTG GTG CTG CCT AGG ACC TTT CTG CTG CAT CAC GCC	96
Arg Leu Ala His Val Val Leu Pro Arg Thr Phe Leu Leu His His Ala	
20 25 30	
ATA CCC CTG GAG CCC GAG ATC ATC TTT TCC ACC TAC ACC CGG TTC AGC	144

Ile	Pro	Leu	Glu	Pro	Glu	Ile	Ile	Phe	Ser	Thr	Tyr	Thr	Arg	Phe	Ser	
		35					40					45				
CGG	TCG	CCA	GGG	TCA	TCC	CGC	CGG	TTG	GTG	GTG	TGT	GGG	AAA	CGT	GTC	192
Arg	Ser	Pro	Gly	Ser	Ser	Arg	Arg	Leu	Val	Val	Cys	Gly	Lys	Arg	Val	
	50					55					60					
CTG	CCA	GGG	GAG	GAA	AAC	CAA	CTT	GCG	TCT	TCA	CCT	TCT	GGT	TTG	GCG	240
Leu	Pro	Gly	Glu	Glu	Asn	Gln	Leu	Ala	Ser	Ser	Pro	Ser	Gly	Leu	Ala	
65					70				75						80	
CTT	AGC	CTG	CCT	CTG	TTT	TCC	CAC	GAT	GGG	AAC	TTT	CAT	CCA	TTT	GAC	288
Leu	Ser	Leu	Pro	Leu	Phe	Ser	His	Asp	Gly	Asn	Phe	His	Pro	Phe	Asp	
				85					90					95		
ATC	TCG	GTA	CTG	CGC	ATT	TCC	TGC	CCT	GGT	TCT	AAT	CTT	AGT	CTT	ACT	336
Ile	Ser	Val	Leu	Arg	Ile	Ser	Cys	Pro	Gly	Ser	Asn	Leu	Ser	Leu	Thr	
		100						105					110			
GTC	AGA	TTT	CTC	TAT	CTA	TCT	CTG	GTG	GTG	GCT	ATG	GGG	GCG	GGA	CGG	384
Val	Arg	Phe	Leu	Tyr	Leu	Ser	Leu	Val	Val	Ala	Met	Gly	Ala	Gly	Arg	
		115					120					125				
AAT	AAT	GCG	CGG	AGT	CCG	ACC	GTT	GAC	GGG	GTA	TCG	CCG	CCA	GAG	GGC	432
Asn	Asn	Ala	Arg	Ser	Pro	Thr	Val	Asp	Gly	Val	Ser	Pro	Pro	Glu	Gly	
	130					135					140					
GCC	GTA	GCC	CAC	CCT	TTG	GAG	GAA	CTG	CAG	AGG	CTG	GCG	CGT	GCT	ACG	480
Ala	Val	Ala	His	Pro	Leu	Glu	Glu	Leu	Gln	Arg	Leu	Ala	Arg	Ala	Thr	
145					150				155						160	
CCG	GAC	CCG	GCA	CTC	ACC	CGT	GGA	CCG	TTG	CAG	GTC	CTG	ACC	GGC	CTT	528
Pro	Asp	Pro	Ala	Leu	Thr	Arg	Gly	Pro	Leu	Gln	Val	Leu	Thr	Gly	Leu	
				165					170					175		
CTC	CGC	GCA	GGG	TCA	GAC	GGA	GAC	CGC	GCC	ACT	CAC	CAC	ATG	GCG	CTC	576
Leu	Arg	Ala	Gly	Ser	Asp	Gly	Asp	Arg	Ala	Thr	His	His	Met	Ala	Leu	
		180					185						190			
GAG	GCT	CCG	GGA	ACC	GTG	CGT	GGA	GAA	AGC	CTA	GAC	CCG	CCT	GTT	TCA	624
Glu	Ala	Pro	Gly	Thr	Val	Arg	Gly	Glu	Ser	Leu	Asp	Pro	Pro	Val	Ser	
		195					200					205				
CAG	AAG	GGG	CCA	GCG	CGC	ACA	CGC	CAC	AGG	CCA	CCC	CCC	GTG	CGA	CTG	672
Gln	Lys	Gly	Pro	Ala	Arg	Thr	Arg	His	Arg	Pro	Pro	Pro	Val	Arg	Leu	
	210					215					220					
AGC	TTG	AAC	CCC	GTC	AAT	GCC	GAT	GTA	CCC	GCT	ACC	TGG	CGA	GAC	GCC	720
Ser	Phe	Asn	Pro	Val	Asn	Ala	Asp	Val	Pro	Ala	Thr	Trp	Arg	Asp	Ala	
225					230				235						240	
ACT	AAC	GTG	TAC	TCG	GGT	GCT	CCC	TAC	TAT	GTG	TGT	GTT	TAC	GAA	CGC	768
Thr	Asn	Val	Tyr	Ser	Gly	Ala	Pro	Tyr	Tyr	Val	Cys	Val	Tyr	Glu	Arg	
				245				250						255		
GGT	GGC	CGT	CAG	GAA	GAC	GAC	TGG	CTG	CCG	ATA	CCA	CTG	AGC	TTC	CCA	816
Gly	Gly	Arg	Gln	Glu	Asp	Asp	Trp	Leu	Pro	Ile	Pro	Leu	Ser	Phe	Pro	
			260					265					270			
GAA	GAG	CCC	GTG	CCC	CCG	CCA	CCG	GGC	TTA	GTG	TTC	ATG	GAC	GAC	TTG	864
Glu	Glu	Pro	Val	Pro	Pro	Pro	Pro	Gly	Leu	Val	Phe	Met	Asp	Asp	Leu	
		275					280					285				
TTC	ATT	AAC	ACG	AAG	CAG	TGC	GAC	TTT	GTG	GAC	ACG	CTA	GAG	GCC	GCC	912
Phe	Ile	Asn	Thr	Lys	Gln	Cys	Asp	Phe	Val	Asp	Thr	Leu	Glu	Ala	Ala	
	290					295					300					

(2) INFORMATION FOR SEQ ID NO:33:

(A) LENGTH: 454 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met	Asp	Ala	His	Ala	Ile	Asn	Glu	Arg	Tyr	Val	Gly	Pro	Arg	Cys	His
1				5					10					15	
Arg	Leu	Ala	His	Val	Val	Leu	Pro	Arg	Thr	Phe	Leu	Leu	His	His	Ala
			20					25					30		
Ile	Pro	Leu	Glu	Pro	Glu	Ile	Ile	Phe	Ser	Thr	Tyr	Thr	Arg	Phe	Ser
		35					40					45			
Arg	Ser	Pro	Gly	Ser	Ser	Arg	Arg	Leu	Val	Val	Cys	Gly	Lys	Arg	Val
	50					55					60				
Leu	Pro	Gly	Glu	Glu	Asn	Gln	Leu	Ala	Ser	Ser	Pro	Ser	Gly	Leu	Ala
65					70					75					80

Leu	Ser	Leu	Pro	Leu	Phe	Ser	His	Asp	Gly	Asn	Phe	His	Pro	Phe	Asp
				85					90					95	
Ile	Ser	Val	Leu	Arg	Ile	Ser	Cys	Pro	Gly	Ser	Asn	Leu	Ser	Leu	Thr
			100					105					110		
Val	Arg	Phe	Leu	Tyr	Leu	Ser	Leu	Val	Val	Ala	Met	Gly	Ala	Gly	Arg
		115					120					125			
Asn	Asn	Ala	Arg	Ser	Pro	Thr	Val	Asp	Gly	Val	Ser	Pro	Pro	Glu	Gly
	130					135					140				
Ala	Val	Ala	His	Pro	Leu	Glu	Glu	Leu	Gln	Arg	Leu	Ala	Arg	Ala	Thr
145					150					155					160
Pro	Asp	Pro	Ala	Leu	Thr	Arg	Gly	Pro	Leu	Gln	Val	Leu	Thr	Gly	Leu
				165					170					175	
Leu	Arg	Ala	Gly	Ser	Asp	Gly	Asp	Arg	Ala	Thr	His	His	Met	Ala	Leu
			180					185					190		
Glu	Ala	Pro	Gly	Thr	Val	Arg	Gly	Glu	Ser	Leu	Asp	Pro	Pro	Val	Ser
		195					200					205			
Gln	Lys	Gly	Pro	Ala	Arg	Thr	Arg	His	Arg	Pro	Pro	Pro	Val	Arg	Leu
	210					215					220				
Ser	Phe	Asn	Pro	Val	Asn	Ala	Asp	Val	Pro	Ala	Thr	Trp	Arg	Asp	Ala
225					230					235					240
Thr	Asn	Val	Tyr	Ser	Gly	Ala	Pro	Tyr	Tyr	Val	Cys	Val	Tyr	Glu	Arg
				245					250					255	
Gly	Gly	Arg	Gln	Glu	Asp	Asp	Trp	Leu	Pro	Ile	Pro	Leu	Ser	Phe	Pro
			260					265					270		
Glu	Glu	Pro	Val	Pro	Pro	Pro	Pro	Gly	Leu	Val	Phe	Met	Asp	Asp	Leu
		275					280					285			
Phe	Ile	Asn	Thr	Lys	Gln	Cys	Asp	Phe	Val	Asp	Thr	Leu	Glu	Ala	Ala
	290					295					300				
Cys	Arg	Thr	Gln	Gly	Tyr	Thr	Leu	Arg	Gln	Arg	Val	Pro	Val	Ala	Ile
305					310					315					320
Pro	Arg	Asp	Ala	Glu	Ile	Ala	Asp	Ala	Val	Lys	Ser	His	Phe	Leu	Glu
				325					330					335	
Ala	Cys	Leu	Val	Leu	Arg	Gly	Leu	Ala	Ser	Glu	Ala	Ser	Ala	Trp	Ile
			340					345					350		
Arg	Ala	Ala	Thr	Ser	Pro	Pro	Leu	Gly	Arg	His	Ala	Cys	Trp	Met	Asp
		355					360					365			
Val	Leu	Gly	Leu	Trp	Glu	Ser	Arg	Pro	His	Thr	Leu	Gly	Leu	Glu	Leu
	370					375					380				
Arg	Gly	Val	Asn	Cys	Gly	Gly	Thr	Asp	Gly	Asp	Trp	Leu	Glu	Ile	Leu
385					390					395					400
Lys	Gln	Pro	Asp	Val	Gln	Lys	Thr	Val	Ser	Gly	Ser	Leu	Val	Ala	Cys
				405					410					415	
Val	Ile	Val													

239

435

440

445

Ile Arg Gly Arg Tyr Gly
450

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 984 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..984
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ATG	TTT	GCT	TTG	AGC	TCG	CTC	GTG	TCC	GAG	GGT	GAC	CCG	GAG	GTG	ACC	48
Met	Phe	Ala	Leu	Ser	Ser	Leu	Val	Ser	Glu	Gly	Asp	Pro	Glu	Val	Thr	
1				5					10					15		
AGT	AGG	TAC	GTC	AAG	GGC	GTA	CAA	CTT	GCC	CTG	GAC	CTT	AGC	GAG	AAC	96
Ser	Arg	Tyr	Val	Lys	Gly	Val	Gln	Leu	Ala	Leu	Asp	Leu	Ser	Glu	Asn	
			20					25					30			
ACA	CCT	GGA	CAA	TTT	AAG	TTG	ATA	GAA	ACT	CCC	CTG	AAC	AGC	TTC	CTC	144
Thr	Pro	Gly	Gln	Phe	Lys	Leu	Ile	Glu	Thr	Pro	Leu	Asn	Ser	Phe	Leu	
		35					40					45				
TTG	GTT	TCC	AAC	GTG	ATG	CCC	GAG	GTC	CAG	CCA	ATC	TGC	AGT	GGC	CGG	192
Leu	Val	Ser	Asn	Val	Met	Pro	Glu	Val	Gln	Pro	Ile	Cys	Ser	Gly	Arg	
	50					55					60					
CCG	GCC	TTG	CGG	CCA	GAC	TTT	AGT	AAT	CTC	CAC	TTG	CCT	AGA	CTG	GAG	240
Pro	Ala	Leu	Arg	Pro	Asp	Phe	Ser	Asn	Leu	His	Leu	Pro	Arg	Leu	Glu	
	65				70					75					80	
AAG	CTC	CAG	AGA	GTC	CTC	GGG	CAG	GGT	TTC	GGG	GCG	GCG	GGT	GAG	GAA	288
Lys	Leu	Gln	Arg	Val	Leu	Gly	Gln	Gly	Phe	Gly	Ala	Ala	Gly	Glu	Glu	
			85					90						95		
ATC	GCA	CTG	GAC	CCG	TCT	CAC	GTA	GAA	ACA	CAC	GAA	AAG	GGC	CAG	GTG	336
Ile	Ala	Leu	Asp	Pro	Ser	His	Val	Glu	Thr	His	Glu	Lys	Gly	Gln	Val	
			100					105					110			
TTC	TAC	AAC	CAC	TAT	GCT	ACC	GAG	GAG	TGG	ACG	TGG	GCT	TTG	ACT	CTG	384
Phe	Tyr	Asn	His	Tyr	Ala	Thr	Glu	Glu	Trp	Thr	Trp	Ala	Leu	Thr	Leu	
		115					120					125				
AAT	AAG	GAT	GCG	CTC	CTT	CGG	GAG	GCT	GTA	GAT	GGC	CTG	TGT	GAC	CCC	432
Asn	Lys	Asp	Ala	Leu	Leu	Arg	Glu	Ala	Val	Asp	Gly	Leu	Cys	Asp	Pro	
	130					135					140					
GGA	ACT	TGG	AAG	GGT	CTT	CTT	CCT	GAC	GAC	CCC	CTT	CCG	TTG	CTA	TGG	480
Gly	Thr	Trp	Lys	Gly	Leu	Leu	Pro	Asp	Asp	Pro	Leu	Pro	Leu	Leu	Trp	
145					150					155					160	

006290-62720960

240

CTG	CTG	TTC	AAC	GGA	CCC	GCC	TCT	TTT	TGT	CGG	GCC	GAC	TGT	TGC	CTG	528
Leu	Leu	Phe	Asn	Gly	Pro	Ala	Ser	Phe	Cys	Arg	Ala	Asp	Cys	Cys	Leu	
			165						170					175		
TAC	AAG	CAG	CAC	TGC	GGT	TAC	CCG	GGC	CCG	GTG	CTA	CTT	CCA	GGT	CAC	576
Tyr	Lys	Gln	His	Cys	Gly	Tyr	Pro	Gly	Pro	Val	Leu	Leu	Pro	Gly	His	
			180					185					190			
ATG	TAC	GCT	CCC	AAA	CGG	GAT	CTT	TTG	TCG	TTC	GTT	AAT	CAT	GCC	CTG	624
Met	Tyr	Ala	Pro	Lys	Arg	Asp	Leu	Leu	Ser	Phe	Val	Asn	His	Ala	Leu	
		195					200					205				
AAG	TAC	ACC	AAG	TTT	CTA	TAC	GGA	GAT	TTT	TCC	GGG	ACA	TGG	GCG	GCG	672
Lys	Tyr	Thr	Lys	Phe	Leu	Tyr	Gly	Asp	Phe	Ser	Gly	Thr	Trp	Ala	Ala	
	210					215					220					
GCT	TGC	CGC	CCG	CCA	TTC	GCT	ACT	TCT	CGG	ATA	CAA	AGG	GTA	GTG	AGT	720
Ala	Cys	Arg	Pro	Pro	Phe	Ala	Thr	Ser	Arg	Ile	Gln	Arg	Val	Val	Ser	
225					230					235					240	
CAG	ATG	AAA	ATC	ATA	GAT	GCT	TCC	GAC	ACT	TAC	ATT	TCC	CAC	ACC	TGC	768
Gln	Met	Lys	Ile	Ile	Asp	Ala	Ser	Asp	Thr	Tyr	Ile	Ser	His	Thr	Cys	
			245					250					255			
CTC	TTG	TGT	CAC	ATA	TAT	CAG	CAA	AAT	AGC	ATA	ATT	GCG	GGT	CAG	GGG	816
Leu	Leu	Cys	His	Ile	Tyr	Gln	Gln	Asn	Ser	Ile	Ile	Ala	Gly	Gln	Gly	
			260					265					270			
ACC	CAC	GTG	GGT	GGA	ATC	CTA	CTG	TTG	AGT	GGA	AAA	GGG	ACC	CAG	TAT	864
Thr	His	Val	Gly	Gly	Ile	Leu	Leu	Leu	Ser	Gly	Lys	Gly	Thr	Gln	Tyr	
		275					280					285				
ATA	ACA	GGC	AAT	GTT	CAG	ACC	CAA	AGG	TGT	CCA	ACT	ACG	GGC	GAC	TAT	912
Ile	Thr	Gly	Asn	Val	Gln	Thr	Gln	Arg	Cys	Pro	Thr	Thr	Gly	Asp	Tyr	
		290				295					300					
CTA	ATC	ATC	CCA	TCG	TAT	GAC	ATA	CCG	GCG	ATC	ATC	ACC	ATG	ATC	AAG	960
Leu	Ile	Ile	Pro	Ser	Tyr	Asp	Ile	Pro	Ala	Ile	Ile	Thr	Met	Ile	Lys	
305					310					315					320	
GAG	AAT	GGA	CTC	AAC	CAA	CTC	TAA									984
Glu	Asn	Gly	Leu	Asn	Gln	Leu										
				325												

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 327 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met	Phe	Ala	Leu	Ser	Ser	Leu	Val	Ser	Glu	Gly	Asp	Pro	Glu	Val	Thr	
1				5					10					15		
Ser	Arg	Tyr	Val	Lys	Gly	Val	Gln	Leu	Ala	Leu	Asp	Leu	Ser	Glu	Asn	
			20					25					30			
Thr	Pro	Gly	Gln	Phe	Lys	Leu	Ile	Glu	Thr	Pro	Leu	Asn	Ser	Phe	Leu	
		35					40					45				
Leu	Val	Ser	Asn	Val	Met	Pro	Glu	Val	Gln	Pro	Ile	Cys	Ser	Gly	Arg	
	50					55					60					

241

Pro Ala Leu Arg Pro Asp Phe Ser Asn Leu His Leu Pro Arg Leu Glu
 65 70 75 80
 Lys Leu Gln Arg Val Leu Gly Gln Gly Phe Gly Ala Ala Gly Glu Glu
 85 90 95
 Ile Ala Leu Asp Pro Ser His Val Glu Thr His Glu Lys Gly Gln Val
 100 105 110
 Phe Tyr Asn His Tyr Ala Thr Glu Glu Trp Thr Trp Ala Leu Thr Leu
 115 120 125
 Asn Lys Asp Ala Leu Leu Arg Glu Ala Val Asp Gly Leu Cys Asp Pro
 130 135 140
 Gly Thr Trp Lys Gly Leu Leu Pro Asp Asp Pro Leu Pro Leu Leu Trp
 145 150 155 160
 Leu Leu Phe Asn Gly Pro Ala Ser Phe Cys Arg Ala Asp Cys Cys Leu
 165 170 175
 Tyr Lys Gln His Cys Gly Tyr Pro Gly Pro Val Leu Leu Pro Gly His
 180 185 190
 Met Tyr Ala Pro Lys Arg Asp Leu Leu Ser Phe Val Asn His Ala Leu
 195 200 205
 Lys Tyr Thr Lys Phe Leu Tyr Gly Asp Phe Ser Gly Thr Trp Ala Ala
 210 215 220
 Ala Cys Arg Pro Pro Phe Ala Thr Ser Arg Ile Gln Arg Val Val Ser
 225 230 235 240
 Gln Met Lys Ile Ile Asp Ala Ser Asp Thr Tyr Ile Ser His Thr Cys
 245 250 255
 Leu Leu Cys His Ile Tyr Gln Gln Asn Ser Ile Ile Ala Gly Gln Gly
 260 265 270
 Thr His Val Gly Gly Ile Leu Leu Leu Ser Gly Lys Gly Thr Gln Tyr
 275 280 285
 Ile Thr Gly Asn Val Gln Thr Gln Arg Cys Pro Thr Thr Gly Asp Tyr
 290 295 300
 Leu Ile Ile Pro Ser Tyr Asp Ile Pro Ala Ile Ile Thr Met Ile Lys
 305 310 315 320
 Glu Asn Gly Leu Asn Gln Leu
 325

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 330 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

242

GGATCCCTCT GACAACCTTC AGATAAAAAA CGTATATGCC CCCTTTTTTTC AGTGGGACAG 60
CAACACCCAG CTAGCAGTGC TACCCCCATT TTTTAGCCGA AAGGATTCCA CCATTGTGCT 120
CGATCCAAC GGATTTGACC CCGTGTTOCC CATGGTCGTG CCGCAGCAAC TGGGGCACGC 180
TATTCTGCAG CAGCTGTTGG TGTACCACAT CTACTCCAAA ATATCGGCCG GGGCCCCGGA 240
TGATGTAAAT ATGGCGGAAC TTGATCTATA TACCACCAAT GTGTCATTTA TGGGGCGCAC 300
ATATCGTCTG GACGTAGACA ACACGGATCC 330

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 627 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGATCCGCTG GCAGGTGGGC GCGCACCTCG TCGGGTAGCT TGGAGACAAA CAGCTCCAGG 60
CCAGTCCGCG CCGTAGCGCC TGCAGGTGCC TCACCACCGG GGCCGGGTCA TGCATCTGT 120
TTAGTCCGGA GAAGATAGGG CCCTTGGGAA GCCGCTGAAC CAGCTCCAGG GTCTCCAAGA 180
TGCGCACCGG TTGTCGGAGC TGTGCGGATA GAGGTTAGGG TAGGTGTCCG GTCCGTCCGT 240
GGGCTCAAAC CTGCCCAGAC ACACCACTGT CTGCTGGGGG ATCATCCTTC TCAGGGAGAT 300
GCATTCTTTG GAAGTAGTGG TAGAGATGGA GCAGACTGCC AGGGCGTTGC AGGAGTGGTG 360
GCGATGGTGC GCACCGTTTT TAAGAAACCC CCCAGGGTGG GCACTCCCGC TCCCTGCAGC 420
ATCTCGGCCT GCTGTACGTC CTTGGCGAAT ATGCGACGAA ATGGGCTGTG CGCACGGGGT 480
CCCAGGGCCG GTCCGGTGGC ATACAGGCCG GTGAGGGCCC CCTGGGTCTG TCCGCCTGGA 540
AACAGGGTGC TGTGAAACAA CAGGTTGCAA GGCCGCGAAT ACCCCTCTGC ACGCTGCTGT 600
GGACGTGGGT GTATGCTCCG TGGATCC 627

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

006290 62720550

(2) INFORMATION FOR SEQ ID NO:39:

- (ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID/NO:39:

GAAATTACCC	ACGAGATCGC	TTCCCTGCAC	ACCGCACTTG	GCTACTCATC	AGTEATCGCC	60
CCGGCCCCACG	TGGCCGCCAT	AACTACAGAC	ATGGGAGTAC	ATTGTCAGGA	CCTCTTTATG	120
ATTTTCCCAG	GGGACGCGTA	TCAGGACCGC	CAGCTGCATG	ACTATATCAA	AATGAAAGCG	180
GGCGTGCAAA	CCGGCTCACC	GGGAAACAGA	ATGGATCACG	TGGGATACAC	TGCTGGGGTT	240
CCTCGCTGCG	AGAACCTGCC	CGGTTTGAGT	CATGGTCAGC	TGGCAACCTG	CGAGATAATT	300
CCCACGCCGG	TCACATCTGA	CGTTGCCT				328

(2) INFORMATION FOR SEQ ID NO:40:

- (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AACACGTCAT GTGCAGGAGT GACATTGTGC CGCGGAGAAA CTCAGATCGC ATCCCGTAAC	60
CACACTGAGT GGGAAAATCT GCTGGCTATG TTTTCTGTGA TTATCTATGC CTTAGATCAC	120
AACTGTCACC CG	132

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

244

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AGCCGAAAGG ATTCCACCAT TCCGTGTTGT CTACGTCCAG

40

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GAAATTACCC ACGAGATCGC AGGCAACGTC AGATGTGA

38

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AACACGTCAT GTGCAGGAGT GACCGGGTGA CAGTTGTGAT CTAAGG

46

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

245

ACAGGGCTGG TTGCCCAGGG T

21

(2) INFORMATION FOR SEQ ID NO:45:

- ```
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:
```

AGTTGCAAAC CAGACCTCAG

20

**SECRET**